National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

UNIVERSITY OF ALBERTA

STUDIES ON THE TRANSCRIPTION OF THE *traM* GENE OF THE F PLASMID: A MODEL FOR THE ROLE OF TraM IN THE INITIATION OF STRAND TRANSFER

BY

SONYA PENFOLD

©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY.

DEPARTMENT OF BIOLOGICAL SCIENCES

Edmonton, Alberta

Spring, 1995

National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file *Votre référence*

Our file *Notre référence*

THE AUTHOR HAS GRANTED AN IRREVOCABLE NON-EXCLUSIVE LICENCE ALLOWING THE NATIONAL LIBRARY OF CANADA TO REPRODUCE, LOAN, DISTRIBUTE OR SELL COPIES OF HIS/HER THESIS BY ANY MEANS AND IN ANY FORM OR FORMAT, MAKING THIS THESIS AVAILABLE TO INTERESTED PERSONS.

L'AUTEUR A ACCORDE UNE LICENCE IRREVOCABLE ET NON EXCLUSIVE PERMETTANT A LA BIBLIOTHEQUE NATIONALE DU CANADA DE REPRODUIRE, PRETER, DISTRIBUER OU VENDRE DES COPIES DE SA THESE DE QUELQUE MANIERE ET SOUS QUELQUE FORME QUE CE SOIT POUR METTRE DES EXEMPLAIRES DE CETTE THESE A LA DISPOSITION DES PERSONNE INTERESSEES.

THE AUTHOR RETAINS OWNERSHIP OF THE COPYRIGHT IN HIS/HER THESIS. NEITHER THE THESIS NOR SUBSTANTIAL EXTRACTS FROM IT MAY BE PRINTED OR OTHERWISE REPRODUCED WITHOUT HIS/HER PERMISSION.

L'AUTEUR CONSERVE LA PROPRIETE DU DROIT D'AUTEUR QUI PROTEGE SA THESE. NI LA THESE NI DES EXTRAITS SUBSTANTIELS DE CELLE-CI NE DOIVENT ETRE IMPRIMES OU AUTREMENT REPRODUITS SANS SON AUTORISATION.

ISBN 0-612-01744-3

Canada

November 28, 1994

Ms. Sonya Penfold
Dept. of Microbiology
University of Alberta
Edmonton, Alberta
CANADA T6G 2E9

Dear Sonya,

You have my permission to use the paper "The Nature of the *traK4* Mutation in the F Sex Factor of *Escherichia coli*," of which I am one of the authors, as part of your doctoral dissertation.

Sincerely,

Ken Usher

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: **Sonya Penfold**

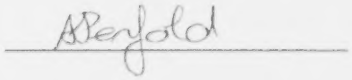
TITLE OF THESIS: **Studies on the Transcription of the *traM* gene of the F Plasmid: A Model for the Role of TraM in the Initiation of Strand Transfer**

DEGREE: **Doctor of Philosophy**

YEAR THIS DEGREE GRANTED: **1995**

Permission is hereby granted to the University of Alberta library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.



Sonya Penfold
3-1616 Beacon Street
Brookline, MA
02146

December 7, 1994

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Studies on the Transcription of the *traM* gene of the F Plasmid: A Model for the Role of TraM in the Initiation of Strand Transfer** submitted by Sonya Penfold in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dr. Laura S. Frost, Supervisor

Dr. William Paranchych

Dr. Brenda Leskiw

Dr. Diane Taylor

Dr. Karen Ippen-Ihler, External Examiner

Date: Nov 30, 1994

To my parents

Alan and Joan

for their continued love and support

Abstract

The expression of genes required for conjugation of the F plasmid of *E. coli* is regulated by a complex cascade of events. Expression of the *tra* operon is inhibited by the FinOP antisense RNA fertility inhibition system, however other, more general control mechanisms also play a role. The F transfer operon was searched for motifs corresponding to those identified for rho-dependent transcriptional terminators and a number of these motifs were identified. In particular the polar nature of a *traK* mutant, *traK4*, was demonstrated to be the result of recognition of a rho-dependent terminator within this gene, resulting in premature termination of transcripts under conditions of reduced translation.

The product of the *traM* gene was previously implicated in events associated with the initiation of DNA transfer from donor to recipient cell and it has been suggested to play a signalling role in this process. TraM is a DNA-binding protein which has been demonstrated to bind to three sites within the origin of transfer (*oriT*).

The *traM* gene lies outside of the major transfer origin and is expressed from its own promoter. An analysis of the regulation of *traM* expression revealed two *traM* promoters, Pm1 and Pm2. Integration host factor, which also binds to *oriT* DNA, was required for maximal expression of *traM*. In addition, the product of the *traY* gene was shown to activate *traM* expression and determine which TraM binding sites on *oriT* are occupied. The amount of TraM protein in F^+ *E. coli* cells correlated well with the relative number of *traM* transcripts that were detected. *traM* expression was also demonstrated to be negatively autoregulated, such that only about 27 molecules of TraM are present in a derepressed cell. In addition, the presence of FinO in cells expressing *traM* resulted in an almost complete repression of *traM*.

A C-terminal TraM mutant which lacked the characteristic acidic tail of F-like TraM proteins was constructed, by deleting the terminal 8 amino acids of the protein. TraMΔ8 was unable to bind to *oriT* DNA *in vitro* or repress *traM* promoters *in vivo*. The protein was demonstrated to have to the same tetrameric conformation as TraM, and a role for the C-terminus of TraM in DNA binding is thus suggested.

The binding characteristics of TraM to *oriT* DNA suggested that a single TraM tetramer binds to each of the two high affinity *oriT* sites. Two protein-DNA complexes were formed when pure protein was incubated with a fragment containing all three TraM binding sites, in contrast to the appearance of only one complex when the protein was present in crude extracts. This was interpreted as a co-operative effect of TraM-*oriT* interactions in the presence of host proteins. In addition, the binding of TraM to *oriT*

DNA required less TraM when pure TraY protein was also present in the reaction, implying that TraY increases the affinity of TraM for *oriT* DNA.

By determining which host and plasmid-encoded proteins play a role in the expression and activity of TraM, we have been able to more clearly map the sequence of events that leads to conjugative transfer of DNA. A model for a role of TraM in the initiation of strand transfer subsequent to origin nicking is proposed.

Acknowledgements

I am grateful to the many people who supported and guided me through the years of completing the work presented in this thesis. Most importantly, my deepest appreciation goes to my supervisor, Dr. Laura Frost, who was always an inspiration to me when I was most frustrated, and could always convince me to see the silver lining. Her continued support and constant enthusiasm for science and life are a lesson to those of us around her. I will leave Edmonton with countless memories of the Rockies, and remember Laura's untiring spirit and body in many hair-raising adventures in the mountains.

Also, to my supervisory committee, Dr. William Paranchych and Dr. Diane Taylor, for their guidance through the years. Special thanks also to Dr. Brenda Leskiw, for her patience and help through many difficult times.

Thanks also the members of the Frost lab, for making the lab a fun place to be: Laura Di Laurenzio, who took me under her wing and did the initial characterization of TraM, Stuart Lee, who reminded us that there was more to life than science, and Tim van Biesen, for showing us the lighter side of life. Then there is also the second generation of Frosties, Karen Anthony, who understood the trials of being a foreign student, and John Simon, who cheerfully helped me whenever he could.

My time in Edmonton would have been much less memorable without my late nights spent with Brenda Blacklock, who always understood the ups and downs. I will always treasure our friendship and marvel at our good fortune at having met.

To my parents, I dedicate this thesis in appreciation of their unselfish love. They were always supportive of my decision to follow my dreams, knowing that they would have to settle for occasional visits. Their constant interest in my life and my hopes despite the miles that separate us, is appreciated.

Finally, to Dwayne, who has always believed in me and helped me believe in myself, my unending love.

Table of Contents

Chapter		Page
1	**Introduction**	
	A. Chromosomal Replication	1
	I. Structure of *oriC*	3
	II. DnaA protein	7
	III. Replication initiation from *oriC*	8
	IV. Regulation of initiation *in vivo*	10
	V. Elongation	13
	VI. Termination	13
	B. Plasmid Replication	14
	I. Origins that use a plasmid-specified initiator protein as well as DnaA	14
	II. Rolling circle replication	20
	C. Conjugation	22
	I. Structure of the F plasmid transfer region	25
	II. Transcription of the *tra* operon	25
	III. Pilus structure and assembly	28
	IV. Regulation of conjugation	29
	V. The *oriT* region	30
	VI. TraY protein	33
	VII. TraI protein	34
	VIII. TraM protein	34
	IX. Donor replacement strand synthesis	36
	X. Recipient second strand synthesis	36
	D. Objectives	37
	E. References	38
2	**The nature of the *traK*4 mutation in the F sex factor of *Escherichia coli***	54
	A. Introduction	55
	B. Results	59
	I. Sequence analysis of the *traK* gene	59
	II. Identification of the *traK*4 and *traK*105 mutations	59
	III. Measurement of termination activity within the *traK* gene	62
	IV. Computer-predicted transcription termination elements (TTEs) in *traK*	62
	V. Analysis of the effectiveness of the TTEs in *traK*	66
	VI. *traK* mRNA secondary structure analysis	71
	C. Discussion	74

D. Materials and Methods 77
I. Bacterial strains, plasmids and media 77
II. Recombinant DNA techniques and reagents 77
III. Assay for transfer efficiency 78
IV. *galK* assay 78
V. Computer-predicted termination sites 78
VI. RNA preparation 78
VII. S1 nuclease analysis 78
VIII. Northern blot analysis 79
IX. Secondary structure analysis 79
X. Genbank accession number 79
E. References 80

3 An analysis of the expression of the *traM* gene of the F sex factor of *Escherichia coli* 84
A. Introduction 85
B. Results 89
I. Construction of pOX*traM*K3 89
II. Mating efficiencies of plasmids used in this study 89
III. Mapping of the 5' ends of *traM* transcripts by primer extension analysis 93
IV. The role of *finO* in *traM* expression 101
V. Analysis of TraM protein production 110
C. Discussion 113
D. Materials and Methods 116
I. Bacterial strains, plasmids and media 116
II. Recombinant DNA techniques and reagents 116
III. Creation of a *traM* insertion mutant 116
IV. Triparental mating to create pOX*traM*K3 119
V. Mating efficiency of pOX*traM*K3 119
VI. Primer extension analysis 119
VII. RNA preparation 120
VIII. Northern blot analysis 120
IX. Analysis of TraM protein production 121
E. References 122

4 The role of the C-terminus of TraM in transfer: Characterization of a 3' deletion mutant of *traM* 127
A. Introduction 128
B. Results 131
I. Construction of a C-terminal mutant, *traM*Δ8 131
II. Immunological detection of TraM and TraMΔ8 131
III. Mating assays 136
IV. FPLC purification of crude extracts 136
V. Gel retardation analysis 142
VI. Primer extension analysis of the autoregulatory 142

capacity of TraMΔ8
C. Discussion 148
D. Materials and Methods 151
I. Bacterial strains, plasmids and media 151
II. Recombinant DNA techniques and reagents 151
III. Construction of a C-terminal deletion mutant (*traM*Δ8) 151
IV. Overexpression of TraM and TraMΔ8 153
V. *In vitro* labelling of overexpressed proteins 153
VI. Mating assays 155
VII. Detection of proteins by immunoblot analysis 155
VIII. Preparation of crude protein extracts for FPLC purification 155
IX. FPLC purification of crude protein extracts 156
X. Gel retardation assays 156
XI. Primer extension analysis 157
E. References 158

5 The binding of TraM to *oriT* DNA 160
A. Introduction 161
B. Results 165
I. The binding of pure TraM to fragment sbmABC 165
II. The binding of TraM in crude extracts to *oriT* DNA 165
III. The effect of TraY on TraM-*oriT* interactions 170
C. Discussion 176
D. Materials and Methods 178
I. Pure protein preparation 178
II. Preparation of crude extracts containing TraM 178
III. Preparation of fragments used in gel retardation assays 178
IV. Gel retardation assays 179
E. References 180

6 General Discussion 182
A. General discussion 183
I. TraM characteristics 183
II. TraM-DNA interactions 183
III. *oriT* secondary structure 184
IV. The role of TraY on *traM* expression 184
V. The role of TraM in transfer 185
VI. Regulation of *traM* by FinO 186
VII. Conjugation versus replication 189
B. Future perspectives 195
C. References 196

List of Tables

Table	Title	Page
2.1	A comparison of *galK* activity observed for wild type, K4 and K105 *traK* genes assayed using a termination assessment vector	63
3.1	Mating efficiencies of plasmid pOX38::Km and its derivatives	92
3.2	Description of bacterial strains used in this study	117
3.3	Description of plasmids used in this study	118
4.1	Mating efficiencies of cells expressing *traM*Δ8 compared with those expressing *traM*	137
4.2	General characteristics of plasmids used to study the effect of *traM*Δ8 on transfer	154

List of Figures

Figure	Title	Page
1.1	Schematic diagram of the structure of *oriC* of *E. coli*	6
1.2	Schematic diagram of the structure of the origin of replication of plasmid P1	17
1.3	Schematic diagram of the structure of the origin of replication of plasmid F	19
1.4	A schematic diagram showing the steps involved in the conjugative transfer of DNA between bacteria	24
1.5	A map of the *tra* operon of the F plasmid of *E. coli*	27
1.6	Features of the origin of transfer (*oriT*) of the F plasmid	32
2.1	A genetic map of the transfer region of the conjugative plasmid F	57
2.2	Nucleotide sequence of the *traK* gene of F	61
2.3	Computer analysis of the *traEK* genes of the F plasmid for potential transcription termination elements	65
2.4	*traK4* transcripts are prematurely terminated	68
2.5	Northern blot analysis of wild type *traK*, *traK4* and *traK105* transcripts	70
2.6	The predicted secondary structure of the wild type *traK*/*traK105* and *traK4* transcripts	73
3.1	The nucleotide sequence of the 1.1 kb *Bgl*II fragment of F	87
3.2	Construction of pOX*traM*K3	91
3.3	Primer extension analysis to detect the 5' ends of *traM* transcripts	95
3.4	A schematic diagram showing the structure of the multicopy plasmids containing cloned fragments of the *oriT* region and *traM*	97
3.5	Autoregulatory activity of TraM	100
3.6	The role of TraY in *traM* expression	103
3.7	FinO decreases *traM* transcription from pOX38::Km	105
3.8	FinO decreases TraM production from F and pOX38::Km	107
3.9	FinO concentration affects TraM expression	109
3.10	A comparison of TraM production from *tra*$^-$ mutants	112
4.1	Amino acid sequence alignment of the 5 TraM alleles	130
4.2	A schematic diagram of the plasmids used to express TraM and TraMΔ8	133
4.3	Western blot analysis of TraM and TraMΔ8 production	135
4.4	TraM and TraMΔ8 are eluted from a gel exclusion column in the same fractions	139
4.5	Calculation of the molecular weight of native TraM and TraMΔ8	141
4.6	TraMΔ8 fails to bind *oriT* DNA	144
4.7	TraMΔ8 is unable to repress *traM* promoters *in vivo*	146

5.1	A comparison of the TraM binding sites to *oriT* in four *traM* alleles	164
5.2	Binding of pure TraM to fragment sbmABC	167
5.3	Binding of TraM in crude protein extracts to fragment sbmABC	169
5.4	Binding of TraM in crude protein extracts to fragment sbmBC and fragment sbmC	172
5.5	The effect of TraY on the binding of TraM to *oriT* DNA	174
6.1	A model for the protein-DNA interactions at F *oriT* in the absence or presence of TraY protein	188
6.2	A summary of the regulation of *traM* expression	191
6.3	A comparison of the F *traM* promoter regions with those of the autoregulated *rep* promoters, *repE* and *repA*	194

Abbreviations

A	adenine
ATP	adenosine triphosphate
bp, kbp	base pair, kilobase pair
BSA	bovine serum albumin
C	cytosine
^{o}C	degrees Centigrade
cm^2	centimetre squared
Da, kDa	Dalton, kiloDalton
dATP	deoxyadenosine triphosphate
Δ	deletion
DNA	deoxyribonucleic acid
DNaseI	deoxyribonuclease I
dNTP	deoxyribonucleotide
DTT	dithiothreitol
EDTA	ethylenediamine tetraacetic acid
F	F fertility factor
fin	fertility inhibition
fmol	femtomole (10^{-15} mol)
g	acceleration of gravity
G	guanine
HEPES	N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
IPTG	isopropyl -β-D-thiogalactopyranoside
IS	insertion sequence
mA	milliAmpere
min	minute
ml	millilitre
µl	microlitre (10^{-6} litre)
ng	nanograms (10^{-9} gram)
ORF	open reading frame
PAG, PAGE	polyacrylamide gel, polyacrylamide gel electrophoresis
pmol	picomole (10^{-12} mole)
RNA	ribonucleic acid
mRNA	messenger RNA
tRNA	transfer RNA
RNase	ribonuclease
SDS	sodium dodecyl sulfate
T	thymine
Tris-HCl	Tris(hydroxymethyl)aminomethane hydrochloride

Chapter 1

Introduction

Ever since the discovery of the complementary double-stranded structure of DNA (Watson and Crick, 1953), the question of how genetic material is replicated and passed from one generation to the next has intrigued researchers. With the development of sophisticated biochemical and molecular biological tools, a detailed picture of the molecular events involved in DNA replication has emerged. However, this picture has become increasingly complex - it has become obvious that different systems have evolved different mechanisms for replicating DNA, and each system involves the interaction of a large number of proteins and other regulatory elements. Although much progress has been made in understanding the events involved in replication, a large number of questions remain unanswered.

Just a few years before the structure of DNA was solved, Lederberg and Tatum (1946) discovered that chromosomal DNA markers could be transferred between certain *E. coli* strains. Although it was soon recognised that the agent responsible for this gene transfer was the F (fertility) factor (Cavalli *et al.*, 1953; Hayes, 1953), an extrachromosomal element, it was only in the 1960's that interest in the properties of conjugative plasmids began to develop. Since then, a large body of information has become available on the physical properties of conjugative plasmids, and more recently, research has focussed on the mechanism of conjugative transfer. As with replication, the picture that is emerging on the mechanism of conjugation is also a complex one, suggesting the involvement of a large number (more than 33 in the case of the F sex factor of *E. coli*) of genes and a surprising number of regulatory elements. Many of the steps involved in the initiation of conjugation are similar to those involved in the initiation of replication, and increasingly, conjugation is being recognised as a variation of a replication event.

Both conjugation and DNA replication can essentially be divided into three steps. In conjugation, the first step involves the formation of a stable mating pair between a donor and recipient cell. Once this stable mating pair has formed, conditions for the initiation of DNA transfer are established and a single strand of plasmid DNA is transferred through the so-called mating bridge to the recipient cell. The final step involves the synthesis of a complementary DNA strand in both the donor and recipient cells, to produce two donor cells, each capable of repeated transfer to suitable recipients. The three essential steps of replication are firstly, initiation, which involves the assembly of the replication complex on the replication origin; secondly, elongation, the synthesis of a second strand of DNA complementary to the replication template; and thirdly, the termination of replication. Although there are obvious common elements to second strand synthesis in both

conjugation and replication, it is the initiation of strand transfer during conjugation and the initiation of replication that bear striking similarities, despite their apparently different biological roles. For this reason, this review will concentrate on describing and comparing the steps involved in the initiation of conjugation and replication. In particular, common steps will be discussed and the associated proteins involved in these steps described. A number of different replication initiation systems have been studied in some detail, and those which appear to bear similarities to conjugation initiation will be emphasized.

A. Chromosomal Replication

Most of our knowledge on the replication of bacterial chromosomes comes from studies on the origin of replication of *E. coli*, *oriC*. However, more recently, alternative origins for chromosomal replication have been described. These are *oriM*, located on two adjacent regions within *oriC* (Asai *et al.*, 1994), and *oriK*(s), located at several sites within the chromosome (Kogoma, 1986). Initiation from these alternative origins is repressed under normal conditions. *oriM* is involved in inducible stable DNA replication (iSDR), which occurs in SOS-induced *E. coli* cells, while *oriK* initiation occurs in *rnhA* mutants which lack RNaseHI activity, and results in constitutive stable DNA replication (cSDR). In all DNA replication initiation systems, the initial step requires that the duplex DNA be melted to allow access of the replication machinery to the template strand. The major difference between *oriC*, *oriM* and *oriK* initiation lies in the mechanism of duplex melting. In *oriC* replication, initial opening is achieved through the formation of a bubble within the duplex by the initiator protein, DnaA. In iSDR, initial duplex opening is achieved through the formation of an intermediate structure (termed a D-loop) of homologous recombination (Asai *et al.*, 1993). The model for cSDR invokes the formation of an R-loop by the displacement of the complementary DNA strand when selected transcripts (which are normally degraded by RNaseHI) hybridize to the template strand (von Meyenburg *et al.*, 1987). In this way, both models for SDR result in the formation of a bubble at the site of replication initiation, allowing entry of the replication apparatus. Since replication by either of these two methods occurs only under abnormal conditions, these will not be discussed further.

Replication of the *E. coli* chromosome occurs bi-directionally and initiates within a specific region, known as *oriC* (the origin of chromosomal replication). Most of our

knowledge about *oriC* has been gained from the use of plasmids containing the *oriC* region, which replicate under the same conditions as the chromosome.

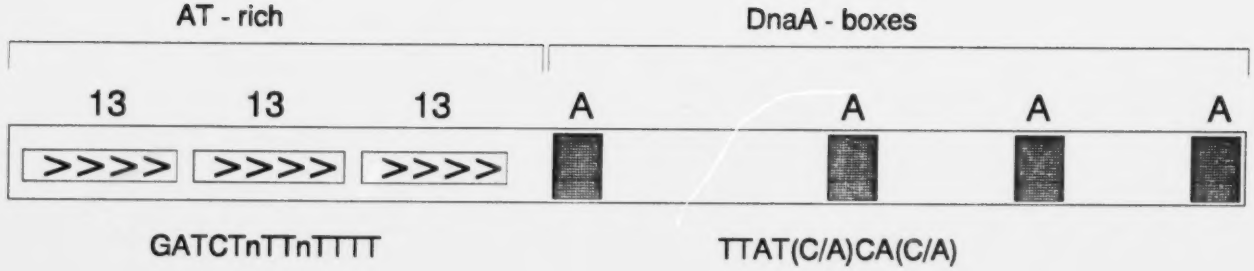
I. Structure of *oriC*

The unique chromosomal origin of replication of *E. coli* has been mapped to about 84 minutes on the genetic map (Bird *et al.*, 1972; Louarn *et al.*, 1974) and is represented schematically in Figure 1.1. The minimal region required for replication is 245 bp (Oka *et al.*, 1980), but maximum efficiency of initiation is dependent on the presence of flanking sequences. The core component of *oriC* consists of an origin recognition element (ORE) and a DNA-unwinding element (DUE). The ORE in *E. coli* consists of a series of four 9 bp repeated sequences (5'-TTATC/ACAC/A), which are binding sites for the initiator protein, DnaA. To the left of the the DnaA binding sites (DnaA-boxes), are three 13-mer sequences (DUEs) which are rich in AT nucleotides. These 13-mers are believed to be involved in duplex-melting subsequent to initial bubble formation by DnaA binding, and it has been demonstrated that the sequence is capable of melting supercoiled DNA, even in the absence of DnaA (Kowalski and Eddy, 1989; Gille and Messer, 1991). In addition to these conserved sequences, another noticeable feature of *oriC* is the large number of GATC sequences. Fourteen of these Dam methylation sites occur within the 300 bp of *oriC*, and although *dam* mutants grow relatively normally, arguing against an essential role for methylation in replication, studies on the replication of plasmids containing *oriC* show a role for methylation in the timing of replication initiation. When hemi-methylated, these *oriC* plasmids bind specifically to cell membrane fractions *in vitro* (Ogden *et al.*, 1988) and are unable to replicate *in vivo* (Russel and Zinder, 1987). Hemi-methylation of GATC sites within *oriC* appears to prevent re-initiation and these sites therefore are important in regulating the rate of replication initiation.

There is a remarkable conservation of sequence within the *oriC* regions of enteric bacteria. Much of this conservation can be attributed to a conservation of the DnaA box regions across a wide variety of bacteria, including a number of Gram positive species (Yoshikawa and Ogasawara, 1991). In fact, homologues of the *dnaA* gene are found in *Pseudomonas putida*, *Pseudomonas aeruginosa*, *Micrococcus luteus*, *Mycoplasma capricolum*, *Proteus mirabilis*, *Streptomyces coelicolor*, *Borrelia burgdorferi*, *Chlamydia trachomatis*, *Spiroplasma citri*, *Rhizobium meliloti*, *Caulobacter crescentus* and *Bacillus subtilis* (reviewed in Skarstad and Boye, 1994). Genome organization both around the *dnaA* gene itself as well as the structure of replication origins share many common

Figure 1.1

Schematic diagram of the structure of *oriC* of *E. coli*. The three 13-mer AT-rich sequences are indicated, with the consensus sequence shown below, where n represents any nucleotide. The four DnaA binding sites are shaded and the DnaA recognition sequence is indicated.


AT - rich
DnaA - boxes
13
13
13
A
A
A
A
>>>>
>>>>
>>>>
GATCTnTTnTTTT
TTAT(C/A)CA(C/A)

features between these disparate bacteria, some of which are believed to have diverged evolutionarily 1.2 billion years ago. This has led to the proposal that an ancestral origin consisted of the *dnaA* gene and a DnaA box region, and subsequent evolution resulted in the duplication of the DnaA boxes, leading to the development of a timing device for replication initiation.

Between the conserved regions of *oriC*, are stretches of non-conserved sequence, which none-the-less seem to play an important role in *oriC* function. These regions appear to serve as spacers which specify a particular conformation of the DNA, as maintenance of the length rather than sequence of these regions is critical for optimal *oriC* function.

II. DnaA protein

The DnaA protein has been purified and shown to be a monomer of 52 kDa that tends to aggregate during purification (Fuller and Kornberg, 1983; Crooke *et al.*, 1993). It is capable of binding both ATP and ADP, but only the ATP-bound form is active in replication. Hydrolysis of ATP from ATP-DnaA is slow, but not required for activity (Sekimizu *et al.*, 1987). Since the rate of exchange of ADP for ATP is also slow, it has been proposed that these nucleotides have an allosteric function in DnaA activity. The rate of exchange of ADP for ATP on DnaA can be increased in the presence of acidic phospholipids (Sekimizu and Kornberg, 1988). In particular, cardiolipin, a cell membrane component, was shown to be able to displace ADP from DnaA and replace it with ATP, thus reconstituting an active form of DnaA for initiation. Furthermore, DnaA could be recovered in membrane fractions of *E.coli* and the purified protein was shown to bind phospholipid vesicles (Sekimizu *et al.*, 1988). Previously, it had been suggested that chromosome replication was dependent on a membrane orientation (Jacob *et al.*, 1963), but no interaction of any replication proteins with the membrane had been demonstrated. The requirement for a membrane location for activation of previously initiation-deficient DnaA, led to the suggestion that this protein may be responsible for attaching the chromosome to the membrane, and thus control initiation.

DnaA is a DNA-binding protein that binds to a characteristic 9 bp sequence, the DnaA box (Fuller *et al.*, 1984). The binding of DnaA to *oriC* is co-operative, with 20-30 monomers of DnaA wrapping the DNA around the outside of the complex (Fuller *et al.*, 1984; Funnel *et al.*, 1987; Crooke *et al.*, 1993). In addition, DnaA binds to its own operator sequence, negatively regulating its own synthesis.

III. Replication initiation from *oriC*

Initiation of replication proceeds through a series of stages. These stages can be differentiated according to the protein and nucleotide requirements of each, as well as the physical changes observed in the DNA by electron microscopy and sensitivity of the DNA to single-stranded specific nucleases.

Initial complex formation involves the recognition of the replication origin by the product of the *dnaA* gene, and subsequent binding of DnaA to the DnaA boxes. From electron microscopic studies of this complex, it appears to be spherical and consists of the DNA wrapped around the outside of the protein aggregate consisting of between 20 and 30 DnaA monomers in a globular structure (Fuller *et al.*, 1984; Funnel *et al.*, 1987; Crooke *et al.*, 1993). Although both the ADP and ATP-bound forms of DnaA are capable of forming this structure, only the complex formed with ATP-bound DnaA is able to promote open complex formation, the next stage of replication initiation (Sekimizu *et al.*, 1987; Crooke *et al.*, 1993).

The second stage of initiation, open complex formation, involves the interaction of DnaA within the protein-DNA complex, with the 13-mer AT-rich regions of *oriC*. Interestingly, by mutating the 13-mer sequences, Bramhill and Kornberg (1988) showed that DnaA specifically recognises the 13-mer sequence, implying that DnaA is capable of recognizing two distinct DNA sequences, the 9-mer DnaA box (as a duplex), and the 13-mer AT-rich sequence, or some part thereof (as either a double- or single-stranded DNA). Additional evidence for the recognition of these 13-mer sequences by DnaA comes from the observation that these 13-mer sequences are conserved in those organisms that show a conservation of the sequence of *dnaA*. Conflicting evidence is provided from work showing that while base substitutions are not tolerated in the rightmost two 13-mers, substitutions in the leftmost 13-mer that retain the AT-richness of the sequence do not affect replication initiation (Hwang and Kornberg, 1992; Asai *et al.*, 1990). As previously mentioned, only the binding of ATP-bound DnaA is capable of melting the 13-mer regions, although not all of the DnaA momoners within the complex need to be ATP-bound. The role of ATP appears to be both allosteric and to provide energy for DNA melting (Sekimizu *et al.*, 1987; Bramhill and Kornberg, 1988). In addition to ATP, a negatively supercoiled template (Baker and Kornberg, 1988), and small amounts of HU protein and integration host factor (IHF) are required (Skarstad *et al.*, 1990). These proteins bind to DNA and play a role in DNA tertiary structure. Specific binding sites for IHF within *oriC* have been demonstrated (Polaczek, 1990) and

although no specific binding site for HU has been found, both proteins have been shown to bend *oriC* DNA (Hodges-Garcia *et al.*, 1989; Polaczek, 1990). Presumably, their role in open complex formation is to bend DNA or alter its conformation in such a way as to promote strand melting.

The previous two steps in replication initiation are both devoted to opening of the duplex by strand separation. The requirement for strand opening is to allow entry of the replication complex into the origin, so that second strand synthesis can start. The final stage of initiation, prepriming complex formation, can be differentiated from the previous one in that it does not have the requirement for high temperature to maintain the open structure. Strand opening by DnaA requires a temperature of 38°C, but this structure can only be maintained after a shift to 16°C, if DnaB and DnaC, two replication enzymes, are also present in the initial, high temperature incubation (Bramhill and Kornberg, 1988). The requirement for a high temperature for DnaA strand-opening reflects the requirement for sufficient energy to open all three 13-mers, while the ability to maintain the open structures in the presence of DnaB and DnaC suggests that these two proteins are responsible for holding the separated strands apart. By probing the initiation complexes with P1, a single stranded-specific nuclease, Funnell *et al.*, (1987) showed that the complex formed at this stage includes the 13-mer regions. This finding is supported by Bramhill and Kornberg (1988) with evidence that the restriction enzyme *Bgl*II no longer cleaves at its recognition sequence located within this region following prepriming complex formation. The large area required to be melted prior to assembly of the replication machinery is not surprising when one considers that the DnaB-DnaC complex is estimated to have a molecular weight of 480 kDa and a Stokes radius of 64Å (Kobori and Kornberg, 1982a).

The mechanism of loading DnaB, a helicase, onto the DNA is not fully understood, but DnaB has a low affinity for both single and double-stranded DNA in the absence of initiation proteins, and appears to require DnaC as an escort (Baker and Wickner, 1991). The DnaB hexamer forms a complex with 6 molecules of DnaC (Wickner and Hurwitz, 1975; Lanka and Schuster, 1983; Kobori and Kornberg, 1982b; Wahle *et al.*, 1989a). The current model proposes that DnaA recognises DnaC in the DnaBC complex and facilitates the loading of DnaB onto the origin (Funnel *et al.*, 1987; Wahle *et al.*, 1989b; Masai *et al.*, 1990). Despite its critical role in the loading of DnaB to the replication complex, DnaC does not remain in the complex. In fact, DnaC inhibits the activity of DnaB, and must be released before DnaB helicase activity is activated (Allen and

Kornberg, 1991). DnaB has a 5'-3' helicase activity that will, in the presence of ATP, single-stranded binding protein (SSB) and DNA gyrase, unwind DNA (Baker *et al.*, 1987; Baker *et al.*, 1986).

IV. Regulation of initiation *in vivo*

The mechanism whereby initiation is co-ordinated with the cell cycle is poorly understood, but a number of factors have been implicated in this process.

There appears to be a requirement for RNA polymerase activity (independent of mRNA or primer synthesis) in replication initiation (Lark, 1972; Messer, 1972). The requirement for transcription in initiation of replication *in vitro* varies depending on temperature, the concentration of HU, or the structure of the template (Baker and Kornberg, 1988; Ogawa *et al.*, 1985). All of the conditions which determine whether initiation is RNA polymerase-dependent affect the structure of *oriC*, implying that the requirement for transcription is to impart the appropriate structure to the origin, so that open complex formation can proceed. It has been shown (Skarstad *et al.*, 1990) that an RNA-DNA hybrid near the origin of replication facilitates the strand separation required for initiation. The ability of this hybrid, located (relatively) far from the site of replication initiation, to activate replication confirms that the requirement for RNA polymerase is not purely to provide a primer for DNA synthesis. An increasing body of evidence suggests that transcription activates DNA replication by influencing the template structure and thus facilitating duplex melting at *oriC*. (Asai *et al.*, 1992; Asai *et al.*, 1990; Ogawa and Okazaki, 1991).

A number of cellular proteins have been implicated in the regulation of initiation. An antagonist to replication initiation acts at the three 13-mers. This protein, IciA, has been shown by gel shift analyses to bind the three 13-mers and thus prevent strand melting by DnaA (Hwang and Kornberg, 1992). IciA has therefore been proposed to be a negative regulator of replication, however, its exact role is not understood, as cells lacking or overproducing the protein did not display a noticeable phenotype (Thöny *et al.*, 1991). The Fis protein, which stimulates inversion in site- specific recombination has been shown to bind *oriC* and exclude DnaA binding in the process (Gille *et al.*, 1991), suggesting a regulatory role for this protein. More recently, it has been demonstrated that the *prs* (phosphoribosylpyrophosphate synthetase) gene product affects the expression of *dnaA*, implying yet another level of control of initation (Sakakibara, 1993).

A critical control in the replication process is the timing of initiation with respect to the cell cycle - the replication of the chromosome must be completed when the mother and daughter cells are ready to part, so that each cell receives a copy of the chromosome. Correspondingly, initiation should be regulated so that initiation does not occur too frequently, resulting in more than one copy of the chromosome per cell. This aspect of regulation of initiation is perhaps the most difficult to understand and the identification of the "molecular clock" that regulates the timing of these events has proven elusive. Nevertheless, there is mounting evidence that suggests that DnaA, aside from its direct role in initiation, plays a role in determining the timing of initiation. Most of this evidence comes from observations of the replication initiation properties of temperature-sensitive (ts) *dnaA* mutants. A *B. subtilis dnaA* (ts) mutant, when shifted from the non-permissive to permissive temperature, showed a number of initiations that correlated well with the amount of DnaA that was in the cell (Moriya *et al.*, 1990). Also, the *E. coli dnaAcos* mutant, which contains three point mutations (Braun *et al.*, 1987), overinitiates, probably as a result of the inability of a factor that normally regulates wild-type DnaA activity to do so on the mutant protein (Skarstad and Boye, 1994). This leads to an accumulation of DNA within the cell and since the concentration of DnaA*cos* within cells is normal, this implies that DnaA has a role in regulating initiation. A third line of evidence for the role of DnaA in regulating initiation comes from a set of amber *dnaA* mutants, which rely on the expression of a suppressor for full-length DnaA production. When this suppressor is inactivated by a shift to the non-permissive temperature, initiation stops rapidly, implying that continued synthesis of DnaA is required (Schaus *et al.*, 1981). In order to investigate whether the concentration of DnaA within cells has an effect on the rate of initiation, a number of groups constructed cells that overexpressed DnaA. Initially, little difference in the concentration of DNA within cells was observed (Churchward *et al.*, 1983; Atlung *et al.*, 1987). However, a clue was provided when the frequencies of specific gene markers were measured by hybridization, and it was determined that *oriC*-proximal markers were dramatically increased. This suggested that the increased DnaA concentration did increase initiation events, but that replication from these initiations was slow or not completed, leading to no observable change in DNA concentration. A more controlled system utilizing a *plac*-controlled *dnaA* gene whose expression could be induced to specific levels by varying concentrations of IPTG was used to extend these findings (Løbner-Olesen *et al.*, 1989). Using flow cytometry, it was shown that the initiation mass (cell mass at the time of initiation /the number of origins to be initiated) and timing of initation was dependent on the concentration of DnaA, within certain limits. This work has been extended further by Atlung and Hansen (1993), who

investigated the effect of increasing DnaA concentrations on the regulation of both the *dnaA* gene and chromosome replication. In addition to finding that the *dnaA* gene was repressed by high levels of DnaA, they found a correlation between DnaA concentration (within certain levels), origin concentration and marker frequency. This elegant study also showed that DnaA overproduction leads to increased replication time as well as affecting the rate of replication at various positions on the chromosome. Close to the origin, initiations were frequent, but replication was slow, resulting in an increased number of origins within the cell, but little increase in the overall amount of DNA. Overall, three different responses were seen to different DnaA concentrations. At moderate levels of DnaA (100-160% of normal levels), the origin concentration and DNA concentration increased proportionally to the levels of initiator protein. Initiation was well synchronized in cells with multiple origins. When the DnaA concentration was increased to up to levels 3 times greater than normal, origin concentration continued to increase, but replication velocity slowed down and became uneven along the chromosome, preventing increases in DNA concentration. Initiation was still fairly synchronous. Further increases in DnaA concentration led to an inability to terminate replication and noticeable asynchrony in initiation. The number of origins per cell decreased relative to the previous stage.

The results discussed above support the initiator titration model proposed by Hansen *et al.*, (1991b) for control of initiation from *oriC* of *E. coli*. This model proposes that DnaA controls initiation by activating the process only when there are sufficient DnaA molecules to bind all the DnaA boxes located around the chromosome. According to this model, overproduction of DnaA leads to initiation of replication which in turn leads to an increase in the number of DnaA boxes in the cell. This titrates out the available DnaA, derepressing the autoregulated *dnaA* promoter. Further increases in DnaA concentration lead to inhibition of replication fork movement, ultimately resulting in a decrease in the concentration of DnaA boxes in the cell. This leads to an increase in the amount of available DnaA and subsequent repression of the *dnaA* promoter. This model relies on constant expression of *dnaA* throughout the cell cycle, a controversial point. Evidence for growth-regulated concentration of DnaA exists (Chiaramello *et al.*, 1989), as does evidence showing that the DnaA concentration is constant and independent of growth rate (Hansen *et al.*, 1991a).

Clearly then, DnaA plays a role both in determining the frequency of initiation as well as in maintaining synchrony of initiation in cells containing multiple origins, but whether it is the molecular clock for DNA replication remains unproven.

V. Elongation

The recruitment of DnaB to the initiation complex provides the critical link between the initiation and elongation phases of replication. Activation of unwinding by DnaB helicase is followed by interaction of DnaB and DnaG primase in the presence of ATP, to activate primer synthesis (Arai and Kornberg, 1979). This complex of DnaB and DnaG will, on a single stranded template, synthesize a 10-60 nucleotide oligoribonucleotide primer. No evidence for a physical interaction between these two proteins exists, so it has been suggested that perhaps DnaB generates a specific secondary structure on the template DNA which is recognised by the primase for primer synthesis (Baker and Wickner, 1991). In this way, multiple primers can be synthesized on the lagging strand template, without a requirement for sequence specificity. These primers are elongated by DNA Pol III holoenzyme, a large multiprotein complex responsible for highly processive DNA synthesis (Wickner *et al.*, 1973; Wickner and Kornberg, 1974).

VI. Termination

Specific termination regions, *ter* sequences, are located diametrically opposite *oriC* and are capable of blocking progress of the replication fork (de Massy *et al.*, 1987; Hill *et al.*, 1987). When two opposing replication forks meet or encounter *ter* sequences, DNA PolI, assisted by RNaseH removes the RNA primers and then fills in the resulting gaps in order to regenerate covalently closed structures. The 5' and 3'ends of the circular chromosome are then juxtaposed for ligation by *E. coli* ligase. In addition, gyrase is required to separate catenanes and introduce negative supercoils into the circular chromosome.

B. Plasmid replication

The similar circular structure of covalently closed circular plasmids to that of the chromosome suggests that similar mechanisms for replication initiation could be used. However, additional regulatory factors apply to plasmid replication, due to the phenomenon of copy number control, whereby plasmids are present at a set number (specific for each plasmid) in each cell. Replication initiation is controlled in such a way that when the number of plasmids in the cell is lower than the set copy number, initiation frequency increases; when the number of plasmids exceeds that specified by the copy number, the frequency of initiations decreases. Another consideration is the feature of plasmid incompatibility, which means that plasmids belonging to the same incompatiblity group are incapable of replication in the same cell. Invariably, the replication of one plasmid will be favoured, leading to a pure line containing only that plasmid type (Novick and Hoppensteadt, 1978). Plasmid survival therefore depends on the ability to be compatible with the chromosome and with other plasmids present in the cell, as well as the ability to replicate independently of the chromosome (at least for plasmids with copy numbers greater than one.) In this section a number of plasmid replication strategies will be discussed, with particular emphasis placed on the those systems which show a similarity to steps involved in conjugation. It is important to note that despite the differences in the systems, all involve the same three basic steps described for initiation of chromosomal replication: recognition of the origin, initial melting of the duplex and assembly of the replication machinery.

I. Origins that use a plasmid-specified initiator protein as well as DnaA

One or more DnaA boxes can be found in the origins of many plasmids eg. ColE1, P1, F, RK2, R6K, pBR322 (reviewed in Skarstad and Boye, 1994). These plasmids all demonstrate a dependence on DnaA for replication, and it is assumed that binding of DnaA to the DnaA boxes is required, or at least stimulates, plasmid replication. However, most of these plasmids also have a requirement for a plasmid-encoded initiator protein, Rep (Zyskind and Smith, 1986; Bramhill and Kornberg, 1988). In addition to the DnaA boxes, the origins of these plasmids contain a characteristic series of repeated sequences, called iterons, which direct the binding of the Rep protein. Other features such as AT-rich sequences containing 13-mers analogous to those at *oriC*, IHF binding sites and GATC methylation sites are also found.

The role of RepA and DnaA in the initiation of replication of plasmid P1 has been studied in some detail (Mukhopadhyay *et al.*, 1993). P1 contains five 19 bp iterons which bind RepA (Abeles, 1986; Sozhamannan and Chattoraj, 1993) and two DnaA boxes (Figure 1.2). Using circular permutation assays, Mukhopadhyay and Chattoraj (1993), showed that upon binding to the RepA binding site, RepA bends DNA. Two dimensional gel electrophoresis determined that when bound to all 5 iterons on supercoiled DNA, RepA absorbs one positive superhelical turn, indicating that the protein wraps the DNA around itself. However, DNA in this complex was no more sensitive than free DNA to KMnO4, which is specific for pyrimidine bases in unstacked DNA. The conclusion from this work was that RepA binding alone did not lead to strand unwinding, and that replication initiation required additional factors to accomplish duplex melting. This additional factor has subsequently been shown to be DnaA (Mokhopadhyay *et al.*, 1993), which not only stimulates binding by RepA, but also increases the sensivity of the P1 origin to single strand specific nucleases and KMnO4, indicating that DnaA, in conjunction with RepA, causes strand opening. In addition, the binding of RepA to the origin stimulated the binding of DnaA, possibly by altering the conformation of the origin DNA. A similar result has been obtained for plasmids R1 and pSC101, where proteins have been shown to stimulate origin binding activity of DnaA (Masai and Arai, 1987; Stenzel *et al.*, 1991; Ortega-Jimenez *et al.*, 1992). Once strand opening has been achieved, DnaB helicase can be loaded (with the aid of DnaC), and, as with chromosome replication, primase and DNA Pol III perform second strand synthesis.

Replication of the F plasmid is regulated in a similar manner to that of P1. RepE, the initiator protein of F binds to four sites (iterons) within *oriS*, which also contains 2 DnaA boxes, and an AT-rich region containing a 13-mer homologous to those in *oriC* (Figure 1.3). Regulation of initiation is controlled by the competitive binding of RepE to the origin iterons and to *incC*, the incompatiblity locus consisting of five 19 bp repeats. How this competitive binding controls the initiation frequency is not yet fully understood.

A comparison of the genes for plasmid-encoded initiator proteins from a variety of plasmids reveals many similarities. Many, including *repA* and *repE*, are autoregulated. This means that like DnaA, their protein products have two functions: they regulate their own synthesis by binding to sites within their own operators (for a full list see Gammie and Crosa, 1991), as well as binding to the origin of replication to act in initiation. The autoregulatory nature of Rep expression is believed to play a role in controlling replication.

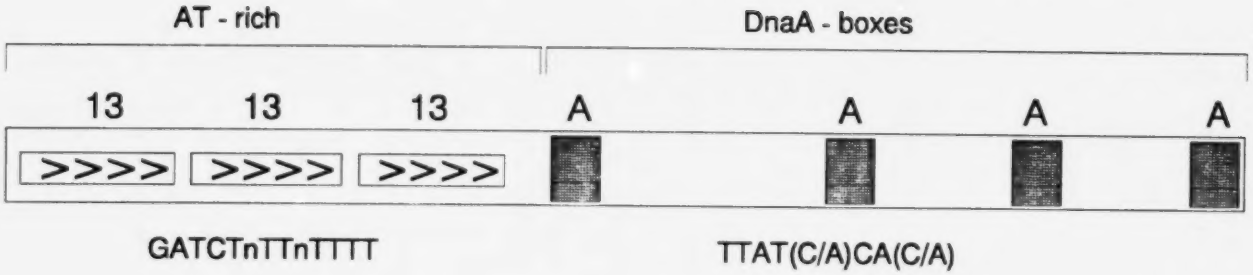
Figure 1.2

Schematic diagram of the structure of the origin of replication of plasmid P1. DnaA binding sites are shaded and the AT-rich region adjacent to the DnaA-boxes is indicated. Five methylation sites contained within the origin are indicated by asterisks, and the five direct repeats to which the plasmid-encoded replication initiation protein RepA binds, are indicated by arrows.

DnaA - boxes
AT - rich
RepA binding sites
*****
Dam methylation sites

Figure 1.3

Schematic diagram of the structure of the origin of replication of plasmid F. DnaA binding sites are shaded and the AT-rich region containing a single site of methylation is indicated. Four direct repeats to which the plasmid-encoded replication initiation protein RepE binds, are indicated by arrows.


AT - rich
DnaA - boxes
13
13
13
A
A
A
A
>>>>
>>>>
>>>>
GATCTnTTnTTTT
TTAT(C/A)CA(C/A)

The presence of primosome assembly sites (*pas*) near the two origins in the RepF1A replicon of F, have led to the suggestion that discontinuous strand synthesis occurs through the action of a primosome. Primosomes are mobile multiprotein complexes which can support priming events at multiple sites and, unlike DnaB and DnaG, are insensitive to the presence of single-stranded binding protein (SSB). The ϕX174 primosome has been well characterised (Arai *et al.*, 1981). Aside from DnaB and DnaG, primosome assembly requires PriA, PriB, PriC, DnaC and DnaT, the prepriming proteins. Assembly occurs at a specific sequence, the *pas*, and after recognition of the site by Pri proteins, DnaB and DnaG associate with the complex to form the primosome. A consensus sequence for the *pas* has been determined (5' GAAGCGG) (Van der Ende *et al.*, 1983), and all *pas* sequences investigated thus far show considerable secondary structure (Greenbaum and Marians, 1985; Arbarzua *et al.*, 1984; Soeller *et al.*, 1984; Marians *et al.*, 1982).

II. Rolling circle replication

Plasmids found in Gram-positive bacteria utilize an entirely different replication mechanism than those of most Gram-negative bacteria. Rolling circle replication begins with the nicking of one strand of the template DNA and extension from the nicked end displaces one of the strands of the duplex. This displaced strand is duplicated after passing through a single-stranded intermediate phase.

The most well studied plasmid that replicates in this fashion is the 4.4 kb plasmid isolated from *Staphylococcus aureus*, pT181. A single replication origin is located within the *repC* gene, which encodes a 38 kDa endonuclease which cleaves at a specific site within the origin (Koepsel and Kahn, 1986). The endonuclease remains bound at the 5' end of the nicked DNA, leaving a free 3'-OH to serve as a primer for DNA synthesis. As the primer is extended, the complementary strand is displaced (Murray *et al.*, 1989). When the synthesis reaches the origin, RepC once again cleaves and then religates the ends of the displaced strand to form a single stranded circle (Iordanescu and Projan, 1988). The synthesis of a complementary strand to this template requires a large, palindromic sequence, *palA* for initiation of second strand synthesis (Gennaro *et al.*, 1989). *repC* expression is regulated by two antisense RNAs, which form RNA-RNA hybrids with the untranslated 5' end of the *repC* mRNA. This structure resembles a Rho-dependent terminator, and causes premature termination of transcription (Kornberg and Baker, 1992).

Despite the variations in replication mechanisms described here, many general steps apply to all systems, whether plasmid or chromosome. In addition, rolling circle replication has been demonstrated to occur in plasmids originally identified in Gram-negative bacteria, eg. pLS1-type plasmids in *Helicobacter pylori* (Kleanthous *et al.*, 1991). For theta-type replication seen in Gram-negative plasmids and the *E. coli* chromosome, initiation of replication always involves the recognition of the origin, the melting of the duplex and subsequent loading of the replication machinery. Different plasmids have different requirements and therefore specify individual replication initiator proteins. Also, since plasmids require the ability to correct copy number, they have additional control systems that not only maintain correct copy number, but allow replication to occur independently of chromosomal replication.

C. Conjugation

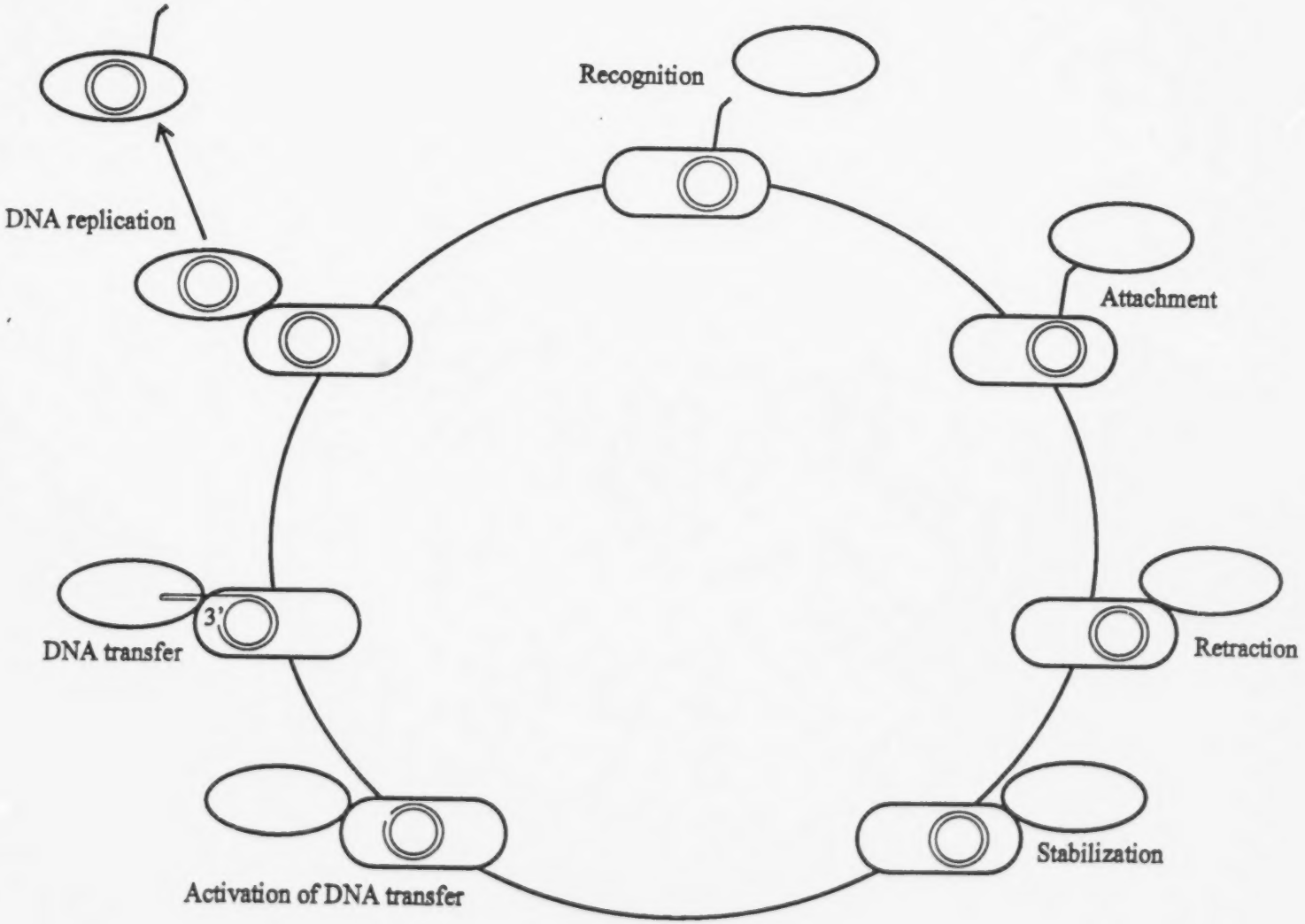
The horizontal transfer of genetic material in bacterial populations represents an enormously powerful mechanism for the spread of genes throughout the environment. The significance of this method of gene transfer has been emphasized recently with the remarkable discovery that conjugation is not restricted to the transfer of DNA between bacteria of similar species, but can occur between bacteria of different genera (Guiney, 1982) and even between kingdoms (Heinemann and Sprague, 1989). While the F sex factor of *E. coli* was the first transmissible element to be discovered (Cavalli *et al.*, 1953; Hayes, 1953), interest in conjugative plasmids was greatly enhanced by the discovery of R plasmids, which carry genes specifying resistance to antibiotics. These plasmids are responsible for the rapid spread of antibiotic resistance genes amongst enterobacteria, resulting in the requirement for new antibiotics to control infection. Other genes found to be transmitted by conjugative plasmids specify toxins, enzymes capable of metabolising complex organic compounds, metal resistance and colonization and pathogenic factors.

Plasmids are classified according to incompatibility (Inc) groups. As previously mentioned, plasmids belonging to the same incompatibility groups are incapable of co-existing within the same cell, probably as a result of competition for replication machinery (Datta, 1975). To date, over 25 different Inc groups have been identified. With the development of recombinant DNA technology, techniques became available to perform sophisicated studies on the molecular events involved in conjugation. The better understood conjugative systems are those of the IncF and IncP groups, eg. F, RP1, RP4 and RK2. What has emerged from research conducted in various laboratories, is that while certain mechanistic aspects of the conjugative process are conserved amongst plasmids, there is a diversity in the biochemical and molecular events required for specific functions. F is the paradigm of conjugative plasmids and there is a large body of information available on F-mediated conjugation (reviewed recently in Frost *et al.*, 1994; Ippen-Ihler and Maneewannakul, 1991; Ippen-Ihler and Skurray, 1993). The events involved in conjugation described in this review, will also focus on transfer of the F plasmid.

Figure 1.4 diagrammatically represents the consecutive steps of the conjugative process. A donor cell, carrying the conjugative plasmid, makes contact with a plasmid-free recipient through the pilus tip. The current model proposes that the pilus is disassembled back into the donor cell, while the tip remains attached to the recipient (Ippen-Ihler and

Figure 1.4

A schematic diagram showing the steps involved in the conjugative transfer of DNA between bacteria. Steps are discussed in the text.


Recognition
Attachment
Retraction
Stabilization
Activation of DNA transfer
DNA transfer
3'
DNA replication

Minkley, 1986). Once the two cells are brought into close contact, the mating pair is stabilized and the (uncharacterised) mating bridge is formed. Mating pair stabilization signals that transfer may begin, and a site- and strand- specific nick is made at *oriT* (the origin of transfer). A single strand of DNA is transferred in the 5'-3' direction to the recipient cell. Once transfer is complete (or perhaps simultaneously), each of the single strands in donor and recipient are replicated, to generate 2 donor cells. In many respects then, conjugation is analogous to a replication event, except that second strand synthesis occurs in two different cells.

I. Structure of the F plasmid transfer region

The entire sequence of the F transfer (*tra*) region is known (Frost *et al.*, 1994). All of the genes required for transfer are located within a 34 kb region. Most of the *tra* genes can be classified into groups according to their role in transfer. Figure 1.5 is a representation of the genes required for transfer, grouped into functional units:

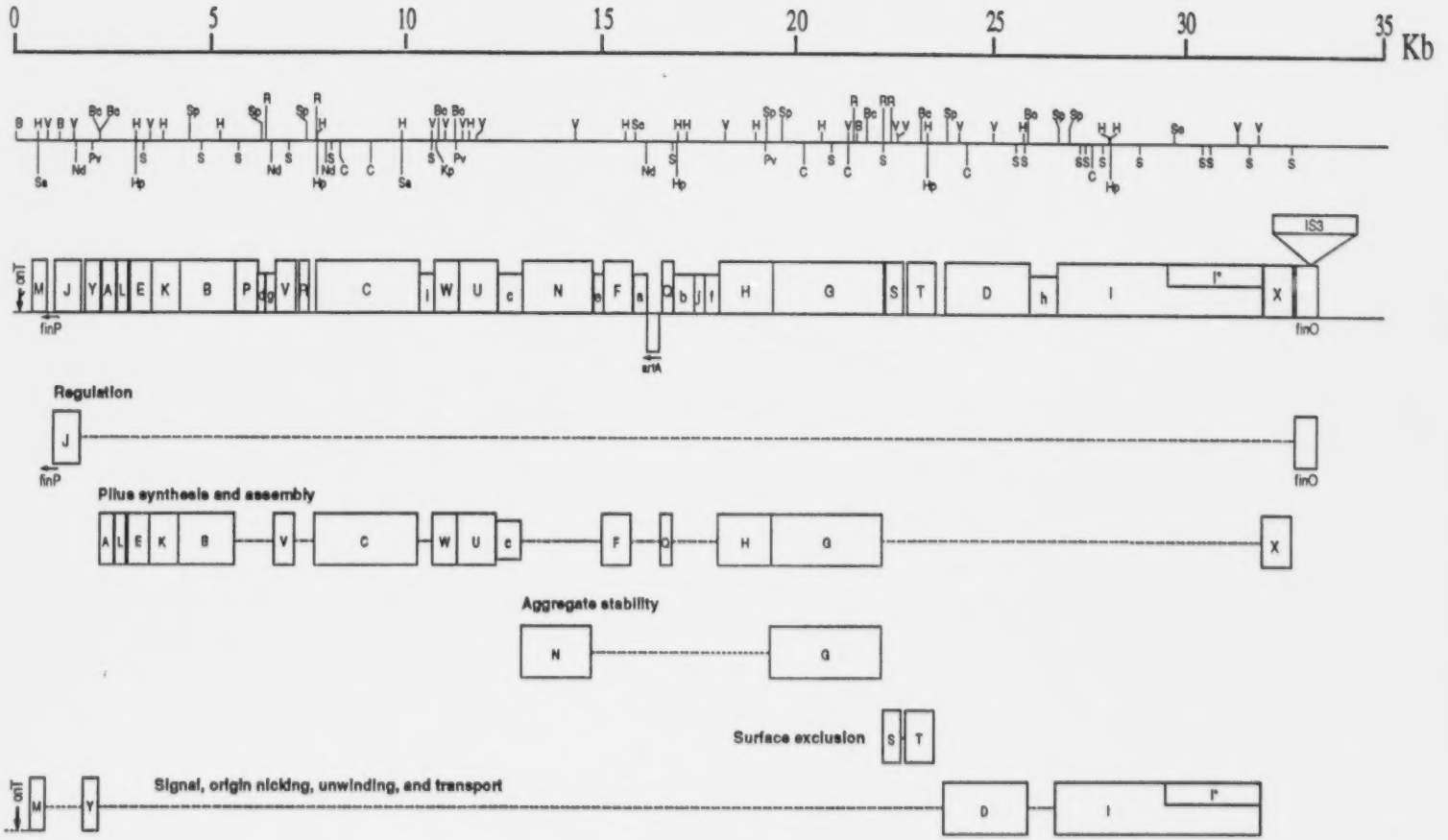
1) Pilus synthesis and assembly: Most of the *tra* genes are devoted to synthesizing pili which mediate attachment to recipient cells.
2) Aggregate Stability: The nature of the stabilization is unknown, but *tra* deficient mutants have been identified which express pili and attach to recipients, but do not form mating aggregates.
3) Surface Exclusion: This prevents transfer to F-containing cells, and is not essential to transfer.
4) Regulation: These genes regulate the expression of transfer functions, determining when transfer occurs.
5) Conjugative DNA metabolism: These genes are involved in transmission of the signal that a stable mating pair / aggregate has formed and that transfer may begin, nicking of DNA and strand transfer. Mutants in these genes are able to express pili and can form stable mating pairs.

II. Transcription of the *tra* operon

With two exceptions, the transfer genes are transcribed in a rightward direction according to Figure 1.5. The major *tra* operon, initiated at Py, the promoter for the *traY* gene, extends over approximately 32 kb, from *traY* to *traI* and perhaps beyond. Although early

Figure 1.5

A map of the *tra* operon of the F plasmid of *E.coli*. The first line indicates the length in kilobases, while the second line shows the position of restriction sites. B=*Bgl*II, Sa=*Sal*I, V=*Eco*RV, Nd=*Nde*I, Bc=*Bcl*I, H=*Hpa*I, S=*Sma*I, Sp=*Sph*I, C=*Cla*I, R=*Eco*RI, Kp=*Kpn*I, arrows=*Hinc*II. The third line indicates the name and position of each gene, where uppercase letters represent *tra* genes and lower case letters represent *trb* genes. All genes are transcribed in a leftward to rightward direction on this map, except for *artA* and *finP*. The remaining lines give the proposed functions of the gene products, according to the current model for conjugation. This figure is from Ippen-Ihler and Maneewannakul (1991).


0
5
10
15
20
25
30
35
Kb
IS3
oriT
finP
finO
traA
Regulation
Pilus synthesis and assembly
Aggregate stability
Surface exclusion
Signal, origin nicking, unwinding, and transport

reports suggested that expression of the *tra* operon was dependent on Py expression (Willetts, 1977), recent evidence has shown that a number of the downstream genes within this operon are expressed independently of Py (reviewed in Frost *et al.*, 1994 and Ippen-Ihler and Maneewannakul, 1991). It is presently unclear how long the transcript expressed from Py is, however, deletion of the Py promoter has been shown to significantly decrease the amount of detectable TraD and TraI protein, indicating that either transcription from Py or a product from the Py transcript is required for efficient expression of these downstream genes (Maneewannakul *et al.*, 1992). In addition to the Py promoter, promoters for the expression of *traM* and *traJ* have been identified by *in vitro* run-off transcription and S1 nuclease experiments (Thompson and Taylor, 1982). The two genes transcribed in the leftward direction are *finP*, an antisense RNA, and *artA*, whose function is unknown.

This thesis concentrates on the genes involved in pilus assembly and conjugal DNA metabolism, so these will be discussed in more detail. Since it is the conjugal DNA metabolism and second strand synthesis that most resemble replication events, particular emphasis will be placed on these events.

III. Pilus structure and assembly

The long flexible pili specified by the F plasmid are composed of a single subunit, pilin, specified by the *traA* gene (Frost *et al.*, 1984). The gene product of *traA* is a 70 amino acid polypeptide which has an acetylated amino terminus, and is processed from a 121 amino acid precursor containing a 51 aa signal sequence. Synthesis of F pili also requires the products of *traQ* and *traX* (reviewed in Ippen-Ihler and Skurray, 1993; Frost *et al.*, 1994), which are required for processing of the precursor and acetylation, respectively. Pilin subunits are arranged in a helix with 5 subunits per turn of the helix (Marvin and Folkhard, 1986). In contrast to the simple structure of the F pilus, a large number of genes are required for assembly. To date, the products of 12 genes have been identified to be required for pilus assembly: *traL*, *traE*, *traK*, *traB*, *traV*, *traC*, *traW*, *trbC*, *traF*, *traH*, *traG* and *traU* (reviewed in Ippen -Ihler and Skurray, 1993 and Frost *et al.*, 1994). The majority of these gene products are predicted to be inner membrane or periplasmic proteins (Frost *et al.*, 1991), supporting their role in pilus assembly. No details on the mechanism of pilus assembly are known.

IV. Regulation of conjugation

The product of the *traJ* gene is the positive regulator of the Py promoter (Willetts, 1977). Regulation of transfer occurs at the level of *traJ* translation. This is achieved through the action of the FinOP (fertility inhibition) antisense RNA system. FinP is a 75 bp antisense RNA, complementary to the untranslated leader sequence and translational start site of *traJ* mRNA (Finlay *et al.*, 1986). FinP RNA has the characteristic secondary structure of antisense RNAs (van Biesen *et al.*, 1993) and is able to bind to the complementary structure within the *traJ* leader sequence. By the classic "kissing reaction" (Tomizawa, 1990), a double stranded structure is formed between the two RNA molecules. This leads to inhibition of *traJ* translation, either by occlusion of the *traJ* translational start site, or by forming a target for cleavage by RNaseIII (van Biesen *et al.*, 1993). Although FinP is expressed constitutively (Dempsey, 1987), it inhibits the production of TraJ only in the presence of another fertility inhibition gene product, FinO, located at the distal end of the *tra* operon. The FinO protein has been demonstrated to extend the half-life of FinP by at least 20-fold (Lee *et al.*, 1992), even in the absence of *traJ* expression. These studies have been extended to show that FinO binds perfectly matched duplex RNA, and therefore binds both FinP and *traJ* mRNA, since each of these transcripts form two stable stem-loop structures (Van Biesen and Frost, 1994). In addition, FinO increases the rate of duplex formation between FinP and *traJ* mRNA 5-fold.

The F plasmid is permanently derepressed for transfer, due to the insertion of an IS*3* element in the *finO* gene (Yoshioka *et al.*, 1987; Cheah and Skurray, 1986). Unlike FinP, FinO is not plasmid specific and expression of FinO from a co-resident plasmid causes a 3-5 fold decrease in detectable *traJ* transcripts as well as repression of the *tra* operon (Lee *et al.*, 1992).

Although expression from the Py promoter is dependent on the *traJ* gene product (Willetts, 1977) and the initiation site has been mapped to within a *Bst*EII site at nt 1789-1795 (Fowler *et al.*, 1983; Mullineaux and Willetts, 1985; Silverman *et al.*, 1991), alternative promoters have been proposed to be active in the absence of the major promoter (Fowler and Thompson, 1986). Weak alternative promoters in this region have also been proposed recently by Silverman *et al.* (1991). Py expression also requires the product of the chromosomally encoded *sfrA* gene (Buxton and Drury, 1983; Lerner and Zinder, 1979; Silverman *et al.*, 1980) and IHF (Gamas *et al.*, 1987; Silverman *et al.*, 1991). Furthermore, evidence for *traJ*-independent expression from Py comes from

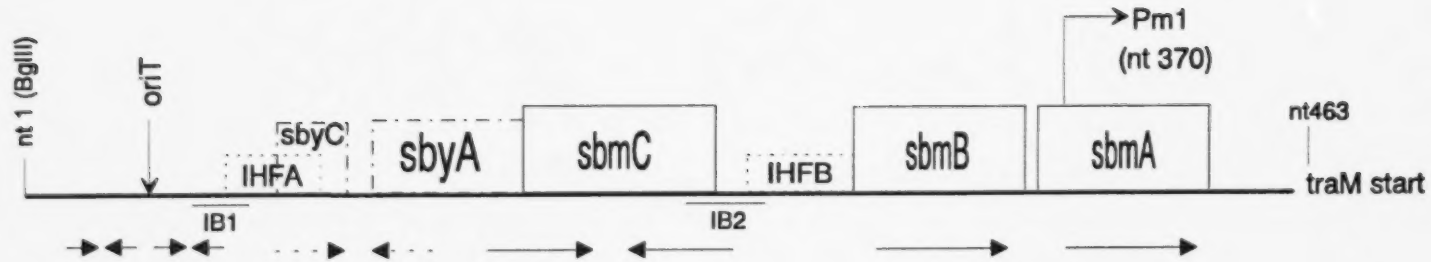
studies on the requirement for *traY* and *traI* in *oriT*-dependent recombination (Carter and Porter, 1991). Control of the Py promoter appears to be complex, depending on the level of supercoiling in the upstream region (Gaudin and Silverman, 1993). In addition, the product of the *traY* gene itself could regulate transfer at some level, as TraY has been shown to bind to a site (*sbyB*) upstream of the Py promoter (Nelson *et al.*, 1993).

V. The *oriT* region

The *oriT* region extends from the *Bgl*II site (arbitrarily labelled nt 1) to the left of the nick site, to the beginning of the *traM* gene (nt 463). Features of *oriT* are shown in Figure 1.6. The nick site is located on the lower strand and has been mapped to just after nt 140 (Thompson *et al.*, 1989; Matson and Morton, 1991: Matson *et al.*, 1993; Reygers *et al.*, 1991). Once nicked, the lower strand is transferred in a 5'-3' direction such that *orf 169* (to the left of *oriT*) is the first sequence transferred. One sequence directed bend, IB1, (caused by runs of repeated homo-polymeric dA•dT bases) (Crothers *et al.*, 1990) is located close to the nick site, while another, IB2, is found approximately 100 bp downstream; together these sites cause a bend centred around nt 245 ((Tsai *et al.*, 1990). In addition, two IHF sites are found within *oriT* (Tsai *et al.*, 1990). IHF is known to bend DNA by more than 140^{o} upon binding (Yang and Nash, 1989), so by binding IHF, *oriT* likely adopts a complicated three-dimensional structure. In addition to bending *oriT*, IHF is required for efficient expression of the *tra* operon and pilus production (Gamas *et al.*, 1987). It is interesting that IHF site A closely resembles IHF binding sites identified in the replication origins of plasmids pSC101 and R6K (γ origin) (Stenzel *et al.*, 1987; Filutowicz and Appelt, 1988), suggesting a possible role for IHF in the replicative stage of conjugation. Between IHF site A and the second internal bend, there is a binding site for the *traY* gene product, *sbyA*. F TraY has been shown to bind to half of an imperfect inverted repeat (Lahue and Matson, 1990; Nelson *et al.*, 1993), recognising a consensus sequence of ATAAA. More recently, a second TraY binding site within *oriT* has been identified (Luo *et al.*, 1994). *sbyC* is located upstream of *sbyA*, in the other half of the imperfect inverted repeat occupied by *sbyA*. This is a low affinity TraY binding site and overlaps IHF site A. Competition gel retardation assays using both IHF and TraY suggest that IHF binds to this DNA segment preferentially to TraY *in vitro*. This study has also shown that TraY bends DNA by 50^{o} upon binding. The centre of TraY-induced bending was located to the centre or left of the centre of *sbyA*. A 17 bp perfect inverted repeat lies adjacent to *sbyA*, overlapping slightly with IB2. Deletion analysis has shown that this region is critical for transfer, but not essential for nicking (Fu *et al.*, 1991). This region

Figure 1.6

Features of the origin of transfer (*oriT*) of the F plasmid. Nucleotide (nt) 1 has arbitrarily been assigned to the *Bgl*II site upstream of the nick site. Sequence-directed bends (IB1 and IB2) are underlined and binding sites identified for TraY, (*sbyA* and *sbyC*) as well as binding sites for IHF (IHFA and IHFB), are indicated. The three TraM binding sites, *sbmA*, *sbmB* and *sbmC* are shown as well as the promoter for the *traM* gene at nt 370. Arrows represent direct and indirect repeats.


nt 1 (BglII)
oriT
sbyC
IHFA
sbyA
sbmC
IHFB
sbmB
sbmA
Pm1
(nt 370)
nt463
traM start
IB1
IB2

has been identified as the low affinity binding site for the product of the *traM* gene (Di Laurenzio *et al.*, 1992), named *sbmC*. Downstream of IHF site B, are two 22bp imperfect direct repeats (Thompson *et al.*, 1984), also shown to be TraM binding sites (Di Laurenzio *et al.*, 1992). A common sequence present in all three of these TraM binding sites is CGGC/TGCG suggesting that this sequence may be a recognition sequence for the TraM protein. It has not been established whether TraM is capable of bending DNA upon binding. Nevertheless, the many features of *oriT*, including intrinsic bends and binding sites for two proteins known to bend DNA (IHF and TraY), as well as direct and inverted repeats, suggests that this region of DNA has a requirement for a specific tertiary structure. The conservation of many of these features in other IncF plasmids (reviewed in Frost *et al.*, 1994), suggests that these features are important for the function of *oriT*.

VI. TraY protein

The product of the *traY* gene is a 15.2 kDa polypeptide (Lahue and Matson, 1990). Interestingly, the F *traY* gene appears to be the result of a gene duplication event, as the N-terminal half and C-terminal half are homologous to each other and to *traY* genes from other *traY* alleles (Inamoto *et al.*, 1988).

Genetic evidence suggests that *traY* is required together with *traI* for the nicking reaction at *oriT* (Everett and Willetts, 1980). However, in *in vitro* nicking assays using either linear or supercoiled templates containing F *oriT*, the presence of purified TraY did not increase the efficiency of nicking by TraI. (Lahue and Matson, 1990; Matson and Morton, 1991). However, under the conditions of the assay, only 50%-70% of the substrate DNA was nicked, suggesting that optimal conditions for nicking had not been established. In a different transfer system, a crude extract containing a TraY fusion polypeptide did contribute to nicking of R100 *oriT* in the presence of overexpressed TraI (Inamoto and Ohtsubo, 1991). This suggests that other, perhaps host encoded products contribute to nicking. It is important to note that neither of these results rule out the possiblity of another function for the TraY protein in transfer.

A possible clue to the function of TraY is given by the finding that TraY is structurally similar to the Arc and Mnt repressors of phage P22 (Bowie and Sauer, 1990), which are known to interact with their operators through their N-terminal domains. This is also the region of *traY* that varies amongst the 4 *traY* alleles. Recent work (Luo *et al.*, 1994) has demonstrated that TraY binds to *sbyA* as a dimer, confirming a previous suggestion

(Nelson *et al.*, 1993) that two TraY dimers mimic the Arc tetramer, which binds DNA by inserting its antiparallel β-sheets into adjacent major grooves at the recognition sequence (Susskind and Youderian, 1983; Breg *et al.*, 1990). Thus, the gene duplication event of F *traY* allows the protein product to dimerize and behave in a manner similar to an Arc tetramer.

VII. TraI protein

In 1983, the large 192 kDa helicase I protein was identified as the product of the *traI* gene (Abdel-Monem *et al.*, 1983). Biochemical characterization of the protein showed it to have single- strand dependent ATPase activity and helicase activity. Extension of these studies showed that helicase I catalyzes a unidirectional (5'-3') and highly processive unwinding reaction that is ATP dependent (Lahue and Matson, 1988). More recently, a second activity of TraI has been identified: that of creating the site- and strand-specific nick at *oriT* (Traxler and Minkley, 1988; Matson and Morton, 1991: Reygers *et al.*, 1991). An analysis of the nicked species produced by TraI suggests that the protein becomes covalently attached to the 5'end of the nicked strand, while the 3' end contains a free OH group available for extension by DNA Polymerase I provided the substrate DNA is treated with SDS or proteinase K (Matson *et al.*, 1993). It is presently unclear whether the nicked species observed in this work represents endonuclease cleavage or an equilibrium between nicking and religation by TraI, as was previously suggested (Everett and Willetts, 1980). TraI then is responsible for creating the nick in *oriT* and unwinding the duplex DNA to generate a single strand for transfer into the recipient cell.

VIII. TraM protein

The *traM* gene of F was mapped to a region outside of the *tra* operon using deletion and complementation analysis (Achtman *et al.*, 1978) and cited as a "sufferance" as it formed an exception to the neatly organized cistrons containing genes of similar function mapped thus far. The ability of *traM* mutants to elaborate pili, form stable mating aggregates and nick *oriT* DNA led to the assignment of the role of transmission of the signal that DNA transfer may begin to the TraM protein (Willetts and Wilkins, 1984). In addition, *traM* has been shown to be required for replacement strand synthesis in the donor cell (Kingsman and Willetts, 1978).

The F TraM protein has been purified (Di Laurenzio *et al.*, 1992) and migrates on SDS-PAGE with an apparent molecular weight of 10,950. The *traM* gene specifies a protein of 127 aa with a predicted molecular weight of 14,507. This anomalous migration on

SDS-PAGE is characteristic of other TraM proteins characterized to date, namely those of pED208 (Di Laurenzio *et al.*, 1991), R1 (Schwab, *et al.*, 1991) and R100 (Dempsey and Fee, 1990). The sequences of the 5 alleles of *traM* are highly homologous, except perhaps for pED208 *traM* (Di Laurenzio, 1992). F TraM is a DNA-binding protein and believed to occur in solution as a tetramer (Di Laurenzio *et al.*, 1992). The protein has been shown to bind to three sites within *oriT* by gel retardation assays and DNaseI footprinting experiments. A high affinity site for TraM, *sbmA* (Figure 1.6), is located within the -10 region of the *traM* promoter, which was identified by run-off transcription experiments using linear templates (Thompson and Taylor, 1982). A site with intermediate affinity for TraM, *sbmB*, is found upstream of *sbmA*, overlapping the -35 region of the *traM* promoter and a 22 bp direct repeat within *oriT*. The site with the lowest affinity for TraM, *sbmC* lies adjacent to the TraY binding site, *sbyA*, and overlaps slightly with internal bend 2. This site also represents the characteristic inverted repeat located in *oriT* that was shown to be essential for transfer (Fu *et al.*, 1991).

The TraM protein is predicted by computer analysis (SURFACEPLOT) to be a cytoplasmic protein, as no transmembrane domains were identified (Di Laurenzio, 1992). Indeed, when inner membrane and cytoplasmic preparations of cells containing the F plasmid are probed for TraM, most of the protein is found in the cytoplasmic fraction, with a small amount detectable in the inner membrane. This corresponds to the results obtained for pED208 TraM and R1 TraM (Di Laurenzio *et al.*, 1991; Schwab *et al.*, 1991).

The location of *sbmA* led to the suggestion that F TraM is autoregulated. This has been shown to be true for the F-like plasmids R1 and R100. Using a chromosomal *traM-lacZ* fusion, Schwab *et al.* (1993), showed that the *traM* promoter is highly active in the absence of a functional TraM protein, and therefore represses its own synthesis. In addition, this autoregulatory function was attributed to the N-terminal region of the protein, where the sequence divergence between the *traM* alleles is greatest. Similarly, using a plasmid-borne *traM-lacZ* fusion, Abo and Ohtsubo (1993) showed that the R100 *traM* gene is repressed by its own gene product. In this work, a distinction was made between the two R100 *traM* promoters identified, as only the most strongly expressed promoter, Pm2, was regulated by TraM. In addition, expression from Pm2 was shown to be repressed to 60% of the level of fully expressed *traM* by IHF. This contrasts to results published earlier for R100, where IHF was shown to stimulate *traM* expression. (Dempsey and Fee, 1990). The reason for this discrepancy is unclear. No other *tra* genes

have been shown to affect the expression of *traM* in either R1 or R100. Early experiments using *lacZ-traM* fusions, (Gaffney *et al.*, 1983) showed that *traJ* was required for F *traM* expression, while no effect of *traJ* on the transcription of F *traM* could be detected when the *traM* promoter was fused to *galK* (Mullineaux and Willetts, 1985).

IX. Donor replacement strand synthesis

Not much information is available on the requirements for synthesis of a replacement strand in the donor cell. DNA polymerase III has been implicated as an *E. coli dnaE* (TS) mutant was capable of transfer, but unable to regenerate a complementary strand in the donor (Kingsman and Willetts, 1978). Whether a primer for DNA synthesis is required is unclear. The previous study implicated a role for RNA polymerase in primer synthesis in *dnaB* (TS) mutants and a second study using the same host showed that an RNA primer was required (Willetts and Wilkins, 1984). However, the use of *dnaB* mutants ruled out the possible contribution of a DnaB -dependent priming mechanism involving DnaG primase, as occurs for *oriC* replication. The location of a putative primosome assembly site within *oriT* (Thompson *et al.*, 1984), suggests that DnaG primase may in fact operate here.

X. Recipient second strand Synthesis

Unlike a number of other transfer systems (Merryweather *et al.*, 1986; Rees and Wilkins, 1989), no F *tra* proteins have been demonstrated to be transferred into the recipient cell (Rees and Wilkins, 1990). Second strand synthesis for F is believed to be carried out by host proteins utilizing a primer generated by RNA polymerase or some DnaB-dependent process (Willetts and Wilkins, 1984). A single-stranded initiation sequence (*ssiE*), thought to allow primer synthesis by primase, has recently been identified on the transferred strand and lies in the leading region to the left of ORF 169 (Nomura *et al.*, 1991).

D. Objectives

Despite the extensive physical characterization of the TraM protein, little knowledge has been gained about its mechanism of action in transfer. Indeed, no evidence for its role in signalling that transfer may begin has substantiated the circumstantial evidence that resulted in its being assigned this role. In her Ph.D. thesis, Laura Di Laurenzio suggested that the occupation of TraM binding sites could be the critical factor required for signalling. The relatively large amounts of protein required to occupy *sbmC*, made this an obvious target for a sensitive signalling system, and it was suggested that perhaps changes in the membrane brought about by pilus depolymerization could be a mechanism for signal transduction. The interaction of TraM with the inner membrane made this a plausible model, as release of TraM from its membrane location could conceivably increase the intracellular pool of available TraM, resulting in binding to *sbmC*. How this could result in transfer, remained a mystery.

A factor to be considered in the interpretation of this data, is that all protein-DNA binding assays were done *in vitro*. The knowledge that supercoiling affects the expression of many genes, together with the obviously complicated tertiary structure of *oriT* and the effects of host encoded proteins (and possibly plasmid encoded proteins), raised the question of whether these results reflected the *in vivo* situation.

In an effort to understand the role of *traM* in transfer, I undertook a study to analyse the expression of the *traM* gene, and identify which factors affected its expression. The tight regulation of a strong promoter, as characterized by R100 Pm2, suggested that at certain times during the conjugative process, large amounts of the protein were required, but at other times the protein was weakly expressed. Also, the location of the *traM* binding sites suggested that this protein plays a role in *oriT* DNA metabolism: *sbmA* and *sbmB* were likely required to control *traM* expression, but the location of *sbmC* close to the nick site and the TraY binding site, together with the knowledge that deletion of this region drastically reduced transfer efficiency, suggested that this may be the site of action of TraM in controlling the initiation of transfer. In addition, mutational analysis of TraM was done to identify functional domains of the protein.

E. References

Abdel-Monem, M., G. Taucher-Scholz, and M. Q. Klinkert. (1983) Identification of *Escherichia coli* DNA helicase I as the *traI* gene product of the F sex factor. *Proc. Natl. Acad. Sci. USA* **80:** 4659-4663.

Abeles. A. L. (1986) P1 plasmid replication: purification and DNA binding activity of the replication protein RepA. *J. Biol. Chem.* **261:** 3548-3555.

Abo, T., and E. Ohtsubo. (1993) Repression of the *traM* gene of plasmid R100 by its own product and integration host factor at one of two promoters. *J. Bacteriol.* **175:** 4466-4474.

Achtman, M., R. A. Skurray, R. Thompson, R. Helmuth, S. Hall, L. Beautin, and A. J. Clark. (1978) Assigment of *tra* cistrons to *Eco*R1 fragments of F sex facter DNA. *J. Bacteriol.* **133:** 1383-1392.

Allen, G. C. Jr., and A. Kornberg. (1991) Fine balance in the regulation of DnaB helicase by DnaC Protein in replication of *Escherichia coli*. *J. Biol. Chem.* **266:** 22096-22101.

Arai, K., and A. Kornberg. (1979) A general priming system employing only *dnaB* replication protein and primase for DNA replication. *Proc. Natl. Acad. Sci. USA* **76:** 4308-4312.

Arai, K., R. Low, J. Kobori, J. Shlomai, and A. Kornberg. (1981) Mechanism of DnaB protein action. V. Association of DnaB protein, protein n', and other prepriming proteins in the primosome of DNA replication. *J. Biol. Chem.* **256:** 5273-5280.

Ararbarzua, P., W. Soeller, and K. J. Marians. (1984) Mutational analysis of primosome assembly sites. I. Distinct classes of mutants in the pBR322 *Escherichia coli* Factor Y DNA assembly sequences. *J.Biol. Chem.* **259:** 14286-14292.

Asai, T., C. P. Chen, T. Nagata, M. Takanami, and M. Imai. (1992) Transcription in vivo within the replication origin of the *Escherichia coli* chromosome: a mechanism for activating initiation of replication. *Mol. Gen. Genet.* **231:** 169-178.

Asai, T., M. Imai, and T. Kogoma. (1994) DNA damage-inducible replication of the *Escherichia coli* chromosome is initiated at separable sites within the minimal *oriC*. *J. Mol. Biol.* **235:** 1459-1469

Asai, T., S. Sommer, A. Bailone, and T. Kogoma. (1993) Homologous recombination-dependent initiation of DNA-replication from DNA damage-inducible origins in *Escherichia coli*. *EMBO J.* **12:** 3287-3295.

Asai, T., M. Takanami, and M. Imai. (1990) The AT richness and *gid* transcription determine the left border of the replication origin of the *E.coli* chromosome. *EMBO J.* **9:** 4065-4072.

Atlung, T., A. Løbner-Olesen, and F .G. Hansen. (1987) Overproduction of DnaA protein stimulates initiation of chromosome replication in *E. coli*. *Mol. Gen. Genet.* **206:** 51-59.

Atlung, T., and F. G. Hansen. (1993) Three distinct chromosome replication states are induced by increasing concentrations of DnaA protein in *Escherichia coli*. *J. Bacteriol* **175:** 6537-6545.

Baker, T. A., Funnell, B. E., and A. Kornberg. (1987) Helicase action of DnaB protein during replication from the *Escherichia coli* chromosomal origin *in vitro*. *J.Biol. Chem.* **262:** 6877-6885.

Baker, T. A., and A. Kornberg. (1988) Transcriptional activation of initiation of replication from the *E.coli* chromosomal origin: an RNA-DNA hybrid near *oriC*. *Cell* **55:** 113-123.

Baker, T. A., K. Sekimizu, B. E. Funnell, and A. Kornberg. (1986) Extensive unwinding of the plasmid template during staged enzymatic initiation of DNA replication from the origin of the *Escherichia coli* chromosome. *Cell* **45:** 53-64.

Baker, T. A., and S. H. Wickner. (1991) Genetics and enzymology of DNA replication in *Escherichia coli*. *Ann. Rev. Biochem.* **60:** 447-477.

Bird, R. E., J. M. Louarn, J. Martuscelli, and L. Caro. (1972) Origin and sequence of chromosomal replication in *Escherichia coli*. *J. Mol. Biol.* **70:** 549-566.

Bowie, J. U., and R. T. Sauer. (1990) TraY proteins of F and related episomes are members of the Arc and Mnt repressor family. *J. Mol. Biol.* **211:** 5-6.

Bramhill, D., and A. Kornberg. (1988) Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the *E.coli* chromosome. *Cell* **52:** 743-755.

Bramhill, D., and A. Kornberg. (1988) A Model for initiation at origins of DNA replication. *Cell* **54:** 915-918.

Braun, R. E., K. O'Day, and A. Wright. (1987) Cloning and characterization of *dnaA* (Cs), a mutation which leads to overinitiation of DNA replication in *Escherichia coli* K12. *J. Bacteriol.* **169:** 3898-3903.

Breg, J. N., J. H. J. van Opheusden, M. J. M. Burgering, R. Boelens, and R. Kaptein.(1990) Structure of Arc repressor in solution: evidence for a family of β-sheet DNA binding proteins. *Nature* (London) **346:** 586-589.

Buxton, R. S., and L. S. Drury. (1983) Cloning and insertional inactivation of the *dye* (*sfrA*) gene, mutation of which affects sex factor F expression and dye sensitivity of *Escherichia coli*. *J. Bacteriol.* **154:** 1309-1314.

Carter, J. R., and R. D. Porter. (1991) *traY* and *traI* are required for *oriT*-enhanced dependent recombination between *lac*-containing plasmids and λ*plac*5. *J. Bacteriol.* **173:** 1027-1034.

Cavalli, L. L., J. Lederberg, and E. M. Lederberg. (1953) An infective factor controlling sex compatibility in *Bacterium coli*. *J. Gen. Microbiol.* **8:** 89-103.

Cheah, K.-C., and R. A. Skurray. (1986) The F plasmid carries an *IS*3 insertion within *finO*. *J. Gen. Microbiol.* **132:** 3269-3275.

Chiaramello, A. E., and J. W. Zyskind. (1989) Expression of *Escherichia coli dnaA* and *mioC* genes as a function of growth rate. *J. Bacteriol.* **171:** 4272-4280.

Churchward, G., P. Holmans, and H. Bremer. (1983) Increased expression of the *dnaA* gene has no effect on DNA replication in a *dnaA+* strain. *Mol. Gen. Genet.* **192:** 506-508.

Crooke, E., R. Thresher, D. S. Hwang, J. Griffith, and A. Kornberg. (1993) Replicatively active complexes of DnaA protein and the *Escherichia coli* chromosomal origin observed in the electron microscope. *J. Mol. Biol.* **233:** 16-24.

Crothers, D. M., T. E. Haran, and J. G. Nadeau. (1990) Instrinsically bent DNA. *J. Biol. Chem.* **265:** 7093-7096.

Datta, N. (1975) Epidemiology and classification of plasmids, p. 9-15. *In* D. Schlessinger (ed.), Microbiology - 1974. American Society for Microbiology, Washington, D.C.

de Massy, B., J. M. Henson, J. Louarn, J. M. Louarn, and J. P. Bouche. (1987) Inhibition of replication forks exiting the terminus region of the *Escherichia coli* chromosome occurs at two loci separated by 5 minutes. *Proc. Natl. Acad. Sci. USA* **84:** 1759-1763.

Dempsey, W.B. 1987. Transcript analysis of the plasmid R100 *traJ* and *finP* genes. *Mol. Gen. Genet.* **209:** 533-544.

Dempsey, W. B., and B. E. Fee. (1990) Integration host factor affects expression of two genes at the conjugal transfer origin of plasmid R100. *Mol. Microbiol.* **4:** 1019-1028.

Di Laurenzio, L. (1992) Protein-DNA interactions at *oriT* regions. Ph. D. Thesis. University of Alberta, Edmonton, Canada.

Di Laurenzio, L., B. B. Finlay, L. S. Frost, and W. Paranchych. (1991) Characterization of the *oriT* region of IncFV plasmid pED208. *Mol. Microbiol.* **5:** 1779-1790.

Di Laurenzio, L., L. S. Frost, and W. Paranchych. (1992) The TraM protein of the conjugative plasmid F binds to the origin of transfer of the F and ColE1 plasmids. *Mol. Microbiol.* **6:** 2951-2959.

Everett, R., and N. Willetts. (1980) Characterization of an *in vivo* system for nicking at the origin of conjugal DNA transfer of the sex factor F. *J. Mol. Biol.* **136:** 129-150.

Filutowicz, M., and K. Appelt. (1988) The integration host factor of *Escherichia coli* binds to multiple sites at plasmid R6K (γ) origin and is essential for replication. *Nucleic Acids Res.* **16:** 3829-3843.

Finlay, B. B., L. S. Frost, W. Paranchych, and N. S. Willetts. (1986) Nucleotide sequences of five IncF plasmid *finP* alleles. *J. Bacteriol.* **167:** 754-757.

Fowler, T., L. Taylor, and R. Thompson. (1983) The control region of the F plasmid transfer operon: DNA sequence of *traJ* and *traY* genes and characterization of the *traY-Z* promoter. *Gene* **26:** 79-89.

Fowler, T., and R. Thompson. (1986) Shadow promoters in the F plasmid transfer operon. *Mol. Gen. Genet.* **202:** 509-511.

Frost, L. S., K. Ippen-Ihler, and R. A. Skurray. (1994) Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol. Rev.* **58:** 162-210.

Frost, L. S., W. Paranchych, and N. S. Willetts. (1984) DNA sequence of the F *traALE* region that includes the gene for F pilin. *J. Bacteriol.* **160:** 395-401.

Frost, L. S., K. Usher, and W. Paranchych. (1991) Computer analysis of the F transfer region. *Plasmid* **25:** 226.

Fu, F. Y.-H., M.-M. Tasi, Y. Luo, and R. C. Deonier. (1991) Deletion analysis of the F plasmid *oriT* locus. *J. Bacteriol.* **173:** 1012-1020.

Fuller, R. S., B. E. Funnel, and A. Kornberg. (1984) The dnaA protein complex with the *E. coli* chromosomal replication origin (*oriC*) and other DNA sites. *Cell* **38:** 889-900.

Fuller, R. S., and A. Kornberg. (1983) Purified DnaA protein in initiation of replication at the *Escherichia coli* chromosomal origin of replication. *Proc. Natl. Acad. Sci. USA.* **80:** 5817-5821.

Funnell, B. E., T. A. Baker and A. Kornberg. (1987) *In vitro* assembly of a prepriming complex at the origin of the *Escherichia coli* chromosome. *J. Biol. Chem.* **262:** 10327-10324.

Gaffney, D., R. Skurray, and N. Willetts. (1983) Regulation of the F conjugation genes studied by hybridization and *tra-lacZ* fusion. *J. Mol. Biol.* **168:** 103-122.

Gamas, P., L. Caro, D. Galas, and M. Chandler. (1987) Expression of F transfer functions depends on the *E. coli* integration host factor. *Mol. Gen. Genet.* **207:** 302-305.

Gammie, A. E., and J. H. Crosa. (1991) Co-operative autoregulation of a replication protein gene. *Mol Microbiol.* **5:** 3015-3023.

Gaudin, G. M., and P. M. Silverman. (1993) Contributions of promoter context and structure to regulated expression of the F plasmid *traY* promoter in *Escherichia coli* K-12. *Mol. Microbiol.* **8:** 335-342.

Gennaro, M. L., S. Iordanescu, R. P. Novick, R. W. Murray, T. R. Steck, and S. A. Khan. (1989) Functional organization of the plasmid pT181 replication origin. *J. Mol. Biol.* **205:** 355-362.

Gille, H., J. B. Egan, A. Roth, and W. Messer. (1991) The FIS protein binds and bends the origin of chromosomal DNA replication, *oriC*, of *Escherichia coli*. *Nucleic Acids Res.* 19: 4167-4172.

Gille, H., and W. Messer. (1991) Localized DNA melting and structural perturbations in the origin of replication, *oriC*, of *Escherichia coli in vitro* and *in vivo*. *EMBO J.* **10:** 1579-1584.

Greenbaum, J. H., and K. J. Marians. (1985) Mutational analysis of primosome assembly sites - evidence for alternative DNA structures. *J. Biol. Chem.* **260:** 12266-12272.

Guiney, D. G. (1982) Host range of conjugation and replication functions of the *Escherichia coli* sex plasmid F' *lac*: comparison with the broad-host-range plasmid RK2. *J. Mol. Biol* **162:** 699-703.

Hansen, F. G., T. Atlung, R. E. Braun, A. Wright, P. Hughes, and M. Kohiyama. (1991a) Initiator (DnaA) protein concentration as a function of growth rate in *Escherichia coli* and *Salmonella typhimurium*. *J. Bacteriol.* **173:** 5149-5199.

Hansen, F. G., B. B. Christensen, and T. Atlung. (1991b) The initiator titration model: computer simulation of chromosome and minichromosome control. *Res Microbiol.* **142:** 161-167.

Hayes, W. (1953) Observations on a transmissible agent determining sexual differentiation in *Bacterium coli*. *J. Gen. Microbiol.* **8:** 72-88.

Heinemann. J. A., and G. F. Sprague Jr. (1989) Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature* (London) **340:** 205-209.

Hill, T. M., J. M Henson, and P. L. Kuempel. (1987) The terminus region of the *Escherichia coli* chromosome contains two separate loci that exhibit polar inhibition of replication. *Proc. Natl. Acad. Sci. USA* **84:** 1745-1758.

Hodges-Garcia, Y., P. J. Hagerman, D. E. Pettijohn. (1989) DNA ring closure mediated by protein HU. *J. Biol. Chem.* **264:** 14621-14623.

Hwang, D. S., and A. Kornberg. (1992) Opposed actions of the regulatory proteins, DnaA and IciA, in opening the replication origin of *Escherichia coli*. *J. Biol. Chem.* **267:** 23087-23091.

Inamoto, S., Y. Yoshioka, and E. Ohtsubo. (1988) Identification and characterization of the products from the *traJ* and *traY* genes of plasmid R100. *J. Bacteriol.* **170:** 2749-2757.

Inamoto, S., Y. Yoshioka, and E. Ohtsubo. (1991) Site- and strand-specific nicking *in vitro* at *oriT* by the TraY-TraI endonuclease of plasmid R100. *J. Biol. Chem.* **266:** 10086-10092.

Iordanescu, S., and S. J. Projan. (1988) Replication termination for staphylococcal plasmids: plasmids pT181 and pC221 cross-react in the termination process. *J. Bacteriol.* **170:** 3427-

Ippen-Ihler, K., and S. Maneewannakul. (1991) Conjugation among enteric bacteria: mating systems dependent on expression of pili, p. 35-69. *In* M. Dworkin (ed.), Microbial cell-cell interactions. American Society for Microbiology, Washington, D.C.

Ippen-Ihler, K., and E. G. Minkley, Jr. (1986) The conjugation system of F, the fertility factor of *Escherichia coli*. *Ann. Rev. Genet.* **20:** 593-624.

Ippen-Ihler, K., and R. A. Skurray. (1993) Genetic organization of transfer-related determinants on the sex factor F and related plasmids, p. 23-52. *In* D. Clewell (ed.), Bacterial conjugation. Plenum Press, New York.

Jacob, F., S. Brenner, and F. Cuzin. (1963) On the regulation of DNA replication in bacteria. *Cold Spring Harbor Symp. Quant. Biol.* **28:** 329-348.

Kingsman, A., and N. Willetts. (1978) The requirements for conjugal DNA synthesis in the donor strain during F*lac* transfer. *J. Mol. Biol.* **122:** 287-300.

Kleanthous, H., C. L. Clayton, and S. Tabaqchali. (1991) Characterization of a plasmid from *Helicobacter pylori* encoding a replication protein common to plasmids in Gram-positive bacteria. *Mol. Microbiol.* **5:** 2377-2390.

Kobori, J. A., and A. Kornberg. (1982a) The *Escherichia coli* DnaC gene product. III. Properties of the DnaB-DnaC protein complex. *J. Biol. Chem.* **257:** 13770-13775.

Kobori, J. A., and A. Kornberg. (1982b) The *Escherichia coli* DnaC gene product. II. Purification, physical properties and role in replication. *J. Biol. Chem.* **257:** 13763-13769.

Koepsel, R .R, and S. A. Kahn. (1986) Static and initiator protein-enhanced bending of DNA at a replication origin. *Science* 233: 1316-18.

Kogoma, T. (1986) RNase H-defective mutants of *Escherichia coli*. *J. Bacteriol.* **166:** 361-363.

Komberg, A., and T. A. Baker, (1992) Plasmids and Organelles, p. 675. *In* DNA Replication (2nd ed.). W. H. Freeman and Company, N.Y.

Kowalski, D., and M. J. Eddy. (1989) The DNA unwinding element: a novel, cis-acting component that facilitates opening of the *Escherichia coli* replication origin. *EMBO J.* **8:** 4335-4344.

Lahue, E. E., and S. W. Matson. (1988) *Escherichia coli* DNA helicase I catalyzes a unidirectional and highly processive unwinding reaction. *J. Biol. Chem.* **263:** 3208-3215.

Lahue, E. E., and S. W. Matson. (1990) Purified *E.coli* F-factor TraY protein binds *oriT*. *J. Bacteriol.* **172:** 1385-1391.

Lanka, E., and H. Schuster. (1983) The *dnaC* protein of *Escherichia coli*. Purification, physical properties and interaction with *dnaB* protein. *Nucleic Acids Res.* **11:** 987-997.

Lark, K. G. (1972) Evidence for the direct involvement of RNA in the initiation of replication in *Escherichia coli* 15T⁻. *J. Mol. Biol.* **64:** 47-60.

Lederberg, J., and E. L. Tatum. (1946) Gene recombination in *Escherichia coli*. *Nature* (London) **158:** 558.

Lee, S. H., L. S. Frost, and W. Paranchych. (1992) FinOP repression of the F plasmid involves extension of the half life of FinP antisense RNA by FinO. *Mol. Gen. Genet.* **235:** 131-139.

Lerner, T., and N. Zinder. (1979) Chromosomal regulation of sexual expression in *Escherichia coli*. *J. Bacteriol.* **137:** 1063-1065.

Løbner-Olesen, A., K. Skarstadt, F. G. Hansen, K. von Meyenburg, and E. Boye. (1989) The DnaA protein determines the initiation mass of *Escherichia coli* K-12. *Cell* **57:** 881-889.

Louarn, J. M., M. Funderburgh, and R. E. Bird. (1974) More precise mapping of the replication origin in *Escherichia coli* K-12 chromosome. *J. Bacteriol.* **120:** 1-5.

Luo, Y., Q. Gao, and R. Deonier. (1994) Mutational and physical analysis of F plasmid *traY* protein binding at *oriT*. *Mol. Microbiol.* **11:** 449-458.

Maneewannakul, S., K. Maneewannakul, and K. Ippen-Ihler. (1992) Sequence alterations affecting F plasmid transfer gene expression: a conjugation system dependent in transcription by the RNA polymerase of phage T7. *Mol. Microbiol.* **6:** 2961-2973.

Marians, K. J., W. Soeller, and S. L. Zipursky. (1982) Maximal limits of the *Escherichia coli* replication factor Y effector site sequences in pBR322 DNA. *J. Biol. Chem.* **257:** 5656-5662.

Marvin, D. A., and W. Folkhard. (1986) Structure of F-pili: reassessment of symmetry. *J. Mol. Biol* **191:** 299-300.

Masai, H., and K. Arai. (1987) RepA and DnaA proteins are required for initiation of R1 plasmid replication *in vitro* and interact with the *oriR* sequence. *Proc. Natl. Acad. Sci. USA* **84:** 4781-4785.

Masai, H., N. Nomura, and K. Arai. (1990) The ABC-primosome. A novel priming system employing DnaA, DnaB, DnaC, and primase on a hairpin containing a DnaA-box sequence. *J. Biol. Chem.* **265:** 15134-15144.

Matson, S. W., and B. S. Morton. (1991) *Escherichia coli* DNA helicase I catalyzes a site- and strand-specific nicking reaction at the F plasmid *oriT*. *J. Biol. Chem.* **265:** 15124-15133.

Matson, S. W., W. C. Nelson, and B. S. Morton. (1993) Characterization of the reaction product of the *oriT* nicking reaction catalyzed by *Escherichia coli* DNA helicase I. *J. Bacteriol.* **175:** 2599-2606.

Merryweather, A., P. T. Barth, and B. M. Wilkins. (1986) Role and specificity of plasmid RP4-encoded DNA primase in bacterial conjuation. *J. Bacteriol.* **167:** 12-17.

Messer, W. (1972) Initiation of deoxyribonucleic acid replication in *Escherichia coli* B/r: chronology of events and transcriptional control of initiation. *J. Bacteriol.* **112:** 7-12.

Moriya, S., K. Kato, H. Yoshikawa, and N. Ogasawara. (1990) Isolation of a *dnaA* mutant of *Bacillus subtilis* defective in initiation of replication: amount of DnaA protein determines cells intiation potential. *EMBO J.* **9:** 2905-2910.

Mukhopadhyay, G., K. M. Carr, J. M. Kaguni, and D. K. Chattoraj. (1993) Open-complex formation by the host initiator, DnaA, at the origin of P1 plasmid replication. *EMBO J.* **12:** 4547-4554.

Mukhopadhyay, G., and D. K. Chattoraj. (1993) Conformation of the origin of P1 plasmid replication: Initiator protein induced wrapping and intrinsic unstacking. *J. Mol. Biol.*. **231:** 19-28.

Mullineaux, P., and N. Willetts. (1985) Promoters in the transfer region of plasmid F. *Basic Life Sci.* **30:** 605-614.

Murray, R. W., R. R. Koepsel, and S. A. Kahn. (1989) Synthesis of single-stranded plasmid pT181 DNA *in vitro*: Initiation and termination of DNA replication. *J. Biol. Chem.* **264:** 1051-1057.

Nelson, W. C., B. S. Morton, E. E. Lahue, and S. W. Matson. (1993) Characterization of the *Escherichia coli* F factor *traY* gene product and its binding sites. *J. Bacteriol.* **175:** 2221-2228.

Novick, R. P., and F. C. Hoppensteadt. (1978) On plasmid incompatibility. *Plasmid* **1:** 421-4434.

Nomura, N., H. Masai, M. Inuzuka, C. Miyazaki, E. Ohtsubo, T. Itoh, S. Sasamoto, M. Matsui, R. Ishizaki, and K. Arai. (1991) Identification of eleven single-strand initiation sequences (*ssi*) for priming of DNA replication in F, R6K, R100 and ColE2 plasmids. *Gene* **108:** 15-22.

Oka, A., K. Sugimoto, M. Takanami, and Y. Hirota. (1980) Replication origin of the *Escherichia coli* K-12 chromosome: The size and structure of the minimum DNA segment carrying the information for autonomous replication. *Mol. Gen. Genet.* **178:** 9-20.

Ogawa, T., T. A. Baker, A. van der Ende, and A. Kornberg. (1985) Initiation of enzymatic replication at the origin of the *Escherichia coli* chromosome: Contributions of RNA polymerase and primase. *Proc. Natl. Acad. Sci. USA* **82:** 3562-3566.

Ogawa, T., and T. Okazaki. (1991) Concurrent transcription from the *gid* and *mioC* promoters activates replication of an *Escherichia coli* minichromosome. *Mol. Gen. Genet.* **230:** 193-200.

Ogden, G. B., M. J. Pratt, M. Schaechter. (1988) The replication origin of the *E. coli* chromosome binds to cell membranes only when hemimethylated. *Cell* **54:** 127-135.

Ortega-Jiménez, S., R. Giraldo-Suárez, M. E. Fernández-Tresguerres, A. Berzal-Herranz, and R. Díaz-Orejas. (1992) DnaA dependent replication of plasmid R1 occurs in the presence of point mutations that disrupt the dnaA box of *oriR*. *Nucleic Acids Res.* **20:** 2547-2551.

Polaczek, P. (1990) Bending of the origin of replication of *E. coli* by binding of IHF at a specific site. *New Biol.* **2:** 265-271.

Rees. C. E. D., and B. M. Wilkins. (1989) Transfer of *tra* proteins into the recipient cell during bacterial conjugation mediated by plasmid ColIb. *J. Bacteriol.* **171:** 3152-3157.

Rees. C. E. D., and B. M. Wilkins. (1990) Protein transfer into the recipient cell during bacterial conjugation: studies with F and RP4. *Mol. Microbiol.* **4:** 1199-1205.

Reygers, U., R. Wessel, H. Muller, and H. Hoffman-Berling. (1991) Endonuclease activity of *Escherichia coli* DNA helicase I directed against the transfer origin of the F factor. *EMBO J.* **10:** 2689-2694.

Russel, D. W., and N. D. Zinder. (1987) Hemimethylation prevents DNA replication in *Escherichia coli*. *Cell* **50:** 1071-1079.

Sakakibara, Y. (1993) Cooperation of the *prs* and *dnaA* gene products for initiation of chromosome replication in *Escherichia coli*. *J. Bacteriol.* **175:** 5559-5565.

Schaus, N., K. O'Day, W. Peters, and A. Wright. (1981) Isolation and characterization of amber mutations in gene *dnaA* of *Escherichia coli* K-12. *J. Bacteriol.* **145:** 904-913.

Schwab, M., H. Gruber, and G. Högenauer. (1991) The TraM protein of plasmid R1 is a DNA-binding protein. *Mol. Microbiol.* **5:** 439-446.

Schwab, M., H. Reisenzien, and G. Högenauer. (1993) TraM of plasmid R1 regulates its own expression. *Mol. Microbiol.* **7:** 795-803.

Sekimizu, K., D. Bramhill, and A. Kornberg. (1987) ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome. *Cell* **50:** 259-265.

Sekimizu, K., and A. Kornberg. (1988) Cardiolipin activation of dnaA protein, the initiator protein of replication in *Escherichia coli*. *J. Biol. Chem.* **263:** 7131-7135.

Sekimizu, K., B. Y., Yung, and A. Kornberg. (1988) The dnaA protein of *Escherichia coli*. Abundance, improved purification and membrane binding. *J. Biol. Chem.* **263:** 7136-7140.

Silverman, P. M., K. Nat, J. McEwen, and R. Birchman. (1980) Selection of *Escherichia coli* K-12 chromosomal mutants that prevent expression of F plasmid functions. *J. Bacteriol.* **143:** 1519-1523.

Silverman, P. M., E. Wickersham, and R. Harris. (1991) Regulation of the F plasmid *traY* promoter in *Escherichia coli* by host and plasmid factors. *J. Mol. Biol.* **218:** 119-128.

Skarstad, K., T. A. Baker and A. Kornberg. (1990) Strand separation required for initiation of replication at the chromosomal origin of *E.coli* is facilitated by a distant RNA-DNA hybrid. *EMBO J.* **9:** 2341-2348.

Skarstad, K., and E. Boye. (1994) The initiator protein DnaA: evolution, properties and function. *Biochem. Biophys. Acta* **1217:** 111-130.

Soeller, W., P. Arbarzua, and K. J. Marians. (1984) Mutational analysis of primosome assembly sites. II. Role of secondary structure in the formation of active sites. *J. Biol. Chem.* **259:** 14293-14300.

Sozhamannan, S., and D. K. Chattoraj. (1993) Heat shock proteins DnaJ, DnaK, and GrpE stimulate P1 plasmid replication by promoting initiator binding to the origin. *J. Bacteriol.* **175:** 3546-3555.

Stenzel, T. T, T. Macallister, and D. Bastia. (1991) Cooperativity at a distance promoted by the combined action of two replication initiator proteins and a DNA bending protein at the replication origin of pSC101. *Genes Dev.* **5:** 1453-1463.

Stenzel, T. T., P. Patel, and D. Bastia. (1987) The integration host factor of *Escherichia coli* binds to bent DNA at the origin of replication of the plasmid pSC101. *Cell* **49:** 709-717.

Susskind, M. M., and P. Youderian. (1983) Bacteriophage P22 antirepressor an its control, p. 347-366. *In* R. W. Hendrix, J. W. Roberts, F. W. Stahl, and R. Weisberg (ed.), Lambda II. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Thompson, T. L., M. B. Centola, and R. C. Deonier. (1989) Location of the nick at *oriT* of the F plasmid. *J. Mol. Biol.* **207:** 505-512.

Thompson, R., and L. Taylor. (1982) Promoter mapping and DNA sequencing of the F plasmid transfer genes *traM* and *traJ*. *Mol. Gen. Genet.* **188:** 513-518.

Thompson, R., L. Taylor, K. Kelly, R. Everett, and N. Willetts. (1984) The F plasmid origin of transfer: DNA sequence of wild type and mutant origins and location of origin-specific nicks. *EMBO J.* **3:** 1175-1180.

Thony, B., D. S. Hwang, L. Fradkin, and A. Kornberg. (1991) *iciA*, an *Escherichia coli* gene encoding a specific inhibitor of chromosomal initiation of replication. *Proc. Natl. Acad. Sci. USA* **88:** 4066-4070.

Tomizawa, J. I. (1990) Control of ColE1 plasmid replication: Intermediates in the binding of RNAI and RNAII. *J. Mol. Biol.* **212:** 683-694.

Traxler, B. A., and E. G. Minkley Jr. (1988) Evidence that DNA helicase I and *oriT* site-specific nicking are both functions of the F TraI protein. *J. Mol. Biol.* **204:** 205-209.

Tsai, M.-M., Y. H. F. Fu, and R. C. Deonier. (1990) Intrinsic bends and integration host factor binding at F plasmid *oriT*. *J. Bacteriol.* **172:** 4603-4609.

van Biesen, T., and L. S. Frost. (1994) The FinO protein of IncF plasmids binds FinP antisense RNA and its target, *traJ* mRNA, and promotes duplex formation. *Mol. Microbiol.* In press.

van Biesen, T., F. Soderbom, E. G. H. Wagner, and L. S. Frost. (1993) Structural and functional analysis of the FinP antisense RNA regulatory system of the F conjugative plasmid. *Mol. Microbiol.* 10: 35-43.

Van der Ende, A., R. Teertstra, H. G. A. M. Van der Avoort and P. J. Weisbeek. (1983) Initiation signals for complementary strand DNA synthesis on single-stranded plasmid DNA. *Nucleic Acids Res.* **11:** 4957-4975.

von Meyenburg, K., E. Boye, K. Skarstadt, L. Koppes, and T. Kogoma. (1987.) Mode of initiation of constitutive stable DNA replication in RNase H-defective mutants of *Escherichia coli* K-12. *J. Bacteriol.* **169:** 2650-2658.

Wahle, E., R. S. Laksen and A. Kornberg. (1989a) The DnaB-DnaC replication protein complex of *Escherichia coli*. I. Formation and properties. *J. Biol. Chem.* **264:** 2463-2468.

Wahle, E., R. S. Laksen and A. Kornberg. (1989b) The DnaB-DnaC replication protein complex of *Escherichia coli*. II. Role of the complex in mobilizing DnaB functions. *J. Biol. Chem.* **264:** 2469-2475.

Watson, J. D., and F. H. C. Crick. (1953) A structure for deoxyribose nucleic acid. *Nature* **171:** 737-738.

Wickner, S. H., and J. Hurwitz. (1975) Interaction of *Escherichia coli* DnaB and DnaC (D) gene products *in vitro*. *Proc. Natl. Acad. Sci. USA* **72:** 921-925.

Wickner, W., and A. Kornberg. (1974) A holoenzyme form of deoxyribonucleic acid polymerase. III. Isolation and properties. *J. Biol. Chem.* **249:** 6244-6249.

Wickner, W., R. Schekman, K. Geider, and A. Kornberg. (1973) A new form of DNA polymerase III and a copolymerase replicate a long, single -stranded primer-template. *Proc. Natl. Acad. Sci. USA* **70:** 1764-1767.

Willetts, N. (1977) The transcriptional control of fertility in F-like plasmids. *J. Mol. Biol.* **112:** 141-148.

Willetts, N. S., and B. Wilkins. (1984) Processing of plasmid DNA during bacterial conjugation. *Microbiol. Rev.* **48:** 24-41.

Yang, C.-C., and H. A. Nash. (1989) The interaction of *E. coli* IHF protein with its specific binding sites. *Cell* **57:** 869-880.

Yoshikawa, H., and N. Ogasawara. (1991) Structure and function of DnaA and the DnaA-box in eubacteria: evolutionary relationships of bacterial replication origins. *Mol. Microbiol.* **5:** 2589-2597.

Yoshioka, Y., H. Ohtsubo, and E. Ohtsubo. (1987) Repressor gene *finO* in plasmids R100 and F: constitutive transfer of plasmid F is caused by insertion of IS*3* into F *finO*. *J. Bacteriol.* **169:** 619-623.

Zyskind, J. W., and D. W. Smith. (1986) The bacterial origin of replication, *oriC*. *Cell* **46:** 489-490.

Chapter 2

The Nature of the *traK4* Mutation in the F Sex Factor of *Escherichia coli**

*A version of this chapter has been published: Penfold, S. S., K. Usher, and L. S. Frost. (1994) *J. Bacteriol.* **176**: 1924-1931.

A. Introduction

The F plasmid is a self-transmissible plasmid, approximately 100 kb in size, with the genes required for transfer located within a 33.4 kb region adjacent to the origin of transfer (for a review of conjugation, see Frost *et al.*, 1994; Ippen-Ihler and Maneewannakul, 1991; Willetts and Skurray, 1987). At least 15 of the 28 genes known to be required for transfer are involved in pilus synthesis and assembly (Frost *et al.*, 1994; Ippen-Ihler and Maneewannakul, 1991). These genes are encoded within the transfer (*tra*) operon which has been estimated to have a maximum length of 31.5 kb (Frost *et al.*, 1994; Willetts, 1977). The *tra* operon is regulated by the TraJ transcriptional activator which positively controls the expression of *tra* genes from the P_Y promoter.

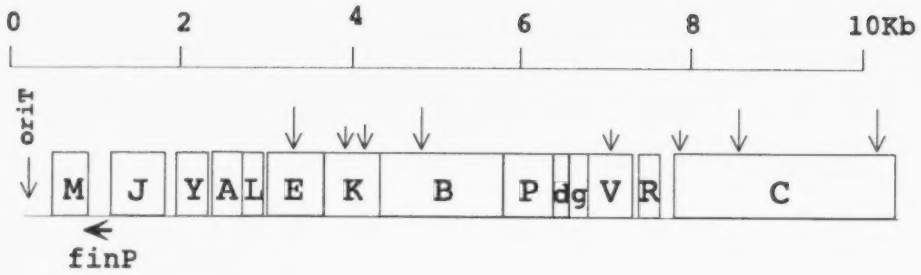
Early experiments suggested that the P_Y transcript extends from *traY* through *traI* (Figure 2.1). Recent work has identified a number of promoters in distal *tra* operon genes and it is presently unclear where the P_Y transcript ends (Ippen-Ihler and Maneewannakul, 1991). Nevertheless, this transcript is predicted to be long, with the next known promoter within the transfer region occurring before the *trbF* gene, nearly 16 kb downstream from the P_Y promoter (Ham *et al.*, 1989).

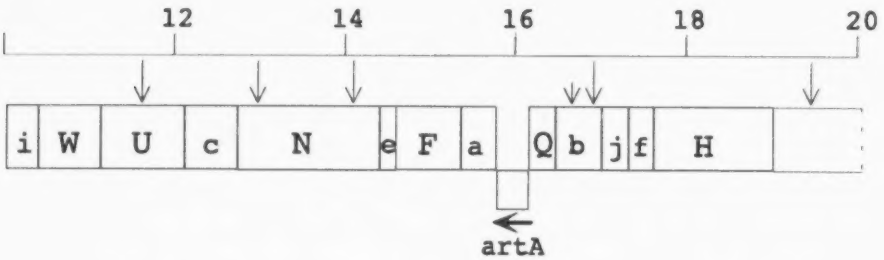
F*lac traK4* and F*lac traK105* (JCFL4 and JCFL105, respectively) are two transfer-deficient mutants of the F plasmid; *tra-4* (later known as F*lac traK4*) was shown to contain an amber-suppressible mutation and was predicted to lie within *traK*, while F*lac traK105* was predicted to contain a frameshift mutation (Achtman *et al.*, 1971; Achtman *et al.*, 1972; Willetts and Achtman, 1972). Complementation tests with F *tra* amber mutants indicated that the *traK*4 mutation was extremely polar and affected expression of genes in the interval from *traK* through *traG* (Ippen-Ihler *et al.*, 1972) (Figure 2.1). Using a series of Hfr strains carrying deletions extending into the transfer region, the position of this polar mutation was confirmed to be within the *traK* gene (Willetts *et al.*, 1976).

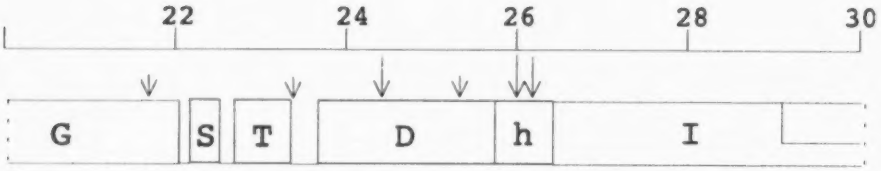
Recently, the salient features of rho-dependent termination in *Escherichia coli* have been characterized (Richardson, 1990; Yager and von Hippel, 1991). A sequence motif common to all rho-dependent terminators characterized thus far, consisting of a region of high cytosine over guanosine content, or C>G-rich bubble, prior to the 3' end of a terminated transcript has been described (Alifano *et al.*, 1991). It was subsequently

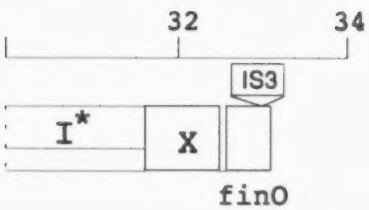
Figure 2.1

A genetic map of the transfer region of the conjugative plasmid F. Included are the positions of C>G bubbles or potential transcription termination elements (TTEs). Arrows above the open reading frames indicate the location of TTEs, where small arrows represent bubbles for which the %C is less than 2 times greater than the %G and large arrows represent bubbles for which the %C is more than 2 times the %G. The scale represents the number of kilobases (kb) from the *Bgl*II site located 141 nt upstream of *oriT*. Upper case letters are *tra* genes while lower case letters are *trb* genes. Two counter transcripts, *finP* and *artA* are indicated. IS*3* represents the insertion element that interrupts the *finO* gene of F.


0
2
4
6
8
10Kb
oriT
M
J
Y
A
L
E
K
B
P
dg
V
R
C
finP


12
14
16
18
20
i
W
U
c
N
e
F
a
Q
b
j
f
H
artA


22
24
26
28
30
G
S
T
D
h
I


32
34
IS3
I*
X
finO

shown that any DNA fragment bearing this consensus motif could activate rho-mediated release of transcripts in the absence of translation, regardless of its physiological role (Rivellini *et al.*, 1991). Richardson (1990, 1991) has proposed that intragenic rho-dependent terminators act to prevent further transcription of long inessential operons during times of stress. Thus, when a transcript containing a latent rho-dependent transcriptional terminator is no longer being translated, the transcriptional termination signal is recognized and transcription ceases.

In this report we present the sequence of the F *traK* gene, identify the positions of the polar *traK4* and the nonpolar *traK105* mutations in the F*lac* plasmid JCFLO and characterize their effect on transcription. Alifano *et al.* (1991) have proposed a consensus sequence for intragenic rho-dependent terminators based on a computer algorithm for DNA sequence analysis. Using a similar program, we predicted the presence of possible termination sites (TTEs) within the transfer operon and used the *traK* mutations to test the validity of these predictions. We have shown that the *traK4* mutation occurs before a predicted termination signal while the *traK105* mutation is located after both predicted TTEs in the *traK* gene. In addition, we verified the position of the 3' ends of the transcripts using S1 nuclease analysis and found that premature termination hastened the rate of degradation of *traK4* transcripts, possibly by subtle alterations in the secondary structure.

B. Results

I. Sequence analysis of the *traK* gene

As part of a project to complete the sequence of the entire F transfer region (Frost *et al.*, 1994), the *traK* gene was cloned into M13mp18 using *Pst*I and *Sma*I restriction sites within the *traE, traK, traB and traP* genes (Frost *et al.*, 1994; Frost *et al.*, 1984). Two *Pst*I fragments of 461 and 2,408 bp extended the sequence presented earlier (Frost *et al.*, 1984) to within the *traP* gene downstream from *traK* and *traB*. A single *Sma*I fragment of 1,498 bp contained the 3' portion of the *traE* gene and extended into the *traB* gene following *traK*. These fragments were sequenced in both directions using oligonucleotide primers to extend the sequence and the portion representing the *traK* gene is shown in Figure 2.2. *traK* is located 1.7 kb downstream from the PY promoter and overlaps with the genes upstream (*traE*) and downstream (*traB*) from it. The *traE* gene overlaps the *traK* sequence by 3 codons while the stop codon for *traK* is part of the initiation codon for *traB*. This pattern of translational coupling is a common feature of the F transfer operon (Frost *et al.*, 1994).

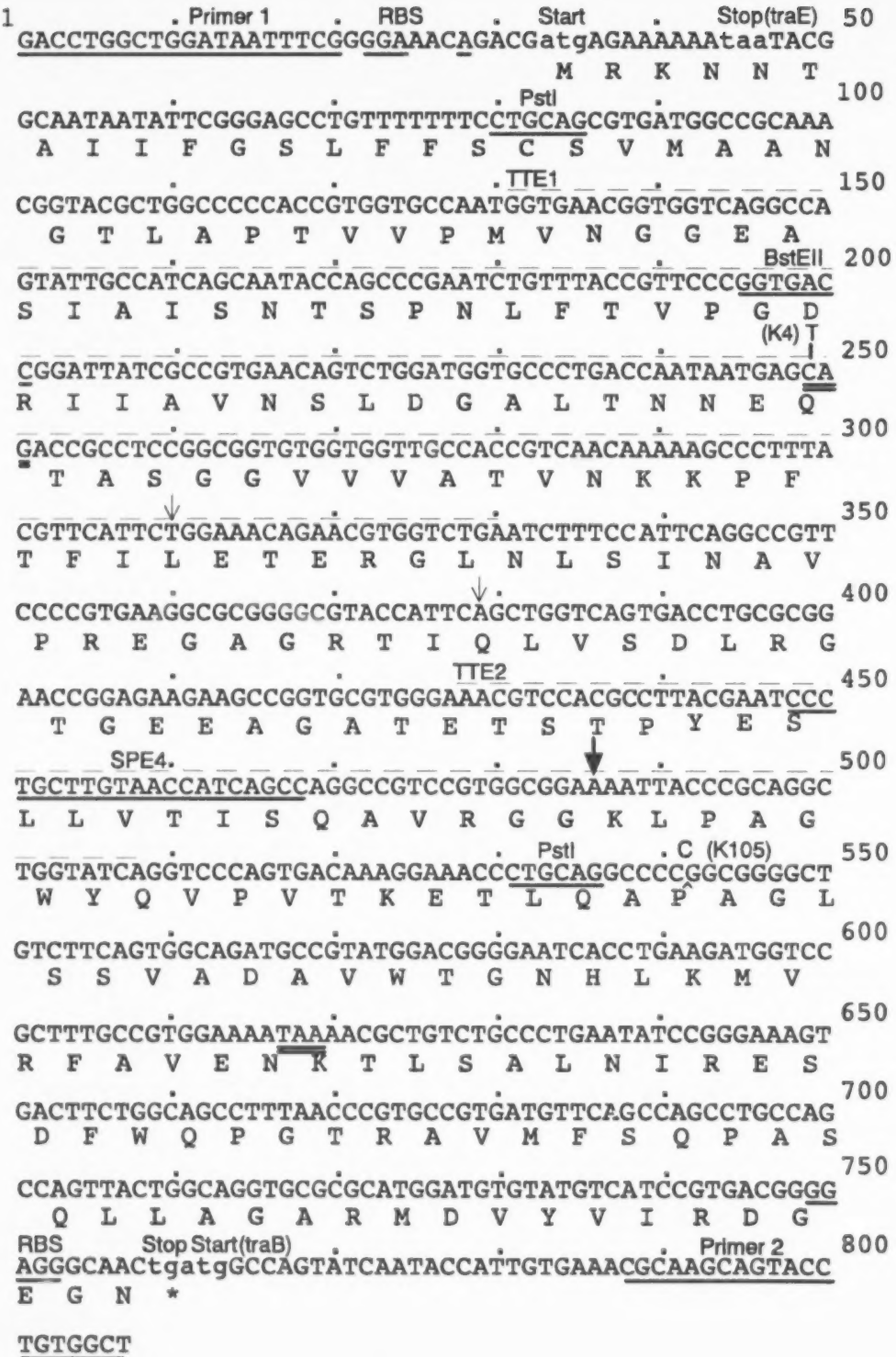
The *traK* gene product is predicted to be a protein of 25,627 Da as determined by the PEPTIDESORT program of PC/GENE. Using the PROSITE and SURFACEPLOT programs, TraK is predicted to be a periplasmic protein of 23,307 Da after cleavage of a signal sequence of 21 amino acids.

II. Identification of the *traK4* and *traK105* mutations

The phenotypes of the F plasmids F*lac traK4* and F*lac traK105* were confirmed by performing mating efficiency assays and comparing the results to those for the wild type F plasmid, JCFL0. The *traK105* mutation could be complemented by wild type *traK* supplied in *trans* from pSPK1, a pUC118 construct containing the wild type *traK* gene, but the *traK4* mutation could not, which agrees with data presented previously (Achtman *et al.*, 1971; Willetts and Achtman, 1972) (data not shown). By using PCR and primers described in Materials and Methods, the *traK* genes from F*lac traK4* and F*lac traK105* were amplified and cloned into pUC118 to give pSPK4 and pSPK5. pSPK4 was found to contain a C-T transition at nucleotide 249 resulting in an amber stop codon at amino acid 73 (Figure 2.2), while the *traK105* mutation resulted from the addition of an extra CG base pair after nucleotide 541, giving a frameshift mutation and termination of translation after codon 195. Four isolates from each amplification and cloning reaction were

Figure 2.2

Nucleotide sequence of the *traK* gene of F. The *traK4* and *traK105* mutations are indicated at nucleotides 249 and 541 respectively and the PCR primers 1 and 2, and oligonucleotide SPE4 are underlined. Lower case letters indicate the overlapping start and stop codons of *traK* and *traE* respectively, as well as the overlapping stop and start codons of *traK* and *traB* respectively. RBS indicates ribosome binding sites for *traK* and *traB*. The stop codons found in *traK4* and *traK105* are double underlined. In addition, the two potential transcription termination elements (TTE) are indicated (—).The open arrows denote the 3' ends of the shorter transcripts found by S1 nuclease analysis of the wild type and K105 *traK* genes, while the filled arrow indicates the 3' end of the *traK4* transcript.


1
Primer 1
RBS
Start
Stop(traE) 50
GACCTGGCTGGATAATTTCGGGGAAACAGACGatgAGAAAAAAtaaTACG
M R K N N T
PstI
100
GCAATAATATTCGGGAGCCTGTTTTTTTCCTGCAGCGTGATGGCCGCAAA
A I I F G S L F F S C S V M A A N
TTE1
150
CGGTACGCTGGCCCCCACCGTGGTGCCAATGGTGAACGGTGGTCAGGCCA
G T L A P T V V P M V N G G E A
BstEII 200
GTATTGCCATCAGCAATACCAGCCCGAATCTGTTTACCGTTCCCGGTGAC
S I A I S N T S P N L F T V P G D
(K4) T
250
CGGATTATCGCCGTGAACAGTCTGGATGGTGCCCTGACCAATAATGAGCA
R I I A V N S L D G A L T N N E Q
300
GACCGCCTCCGGCGGTGTGGTGGTTGCCACCGTCAACAAAAAGCCCTTTA
T A S G G V V V A T V N K K P F
350
CGTTCATTCTGGAAACAGAACGTGGTCTGAATCTTTCCATTCAGGCCGTT
T F I L E T E R G L N L S I N A V
400
CCCCGTGAAGGCGCGGGGCGTACCATTCAGCTGGTCAGTGACCTGCGCGG
P R E G A G R T I Q L V S D L R G
TTE2
450
AACCGGAGAAGAAGCCGGTGCGTGGGAAACGTCCACGCCTTACGAATCCC
T G E E A G A T E T S T P Y E S
SPE4.
500
TGCTTGTAACCATCAGCCAGGCCGTCCGTGGCGGAAAATTACCCGCAGGC
L L V T I S Q A V R G G K L P A G
PstI
C (K105)
550
TGGTATCAGGTCCCAGTGACAAAGGAAACCCTGCAGGCCCCGGCGGGGCT
W Y Q V P V T K E T L Q A P A G L
600
GTCTTCAGTGGCAGATGCCGTATGGACGGGGAATCACCTGAAGATGGTCC
S S V A D A V W T G N H L K M V
650
GCTTTGCCGTGGAAAATAAAACGCTGTCTGCCCTGAATATCCGGGAAAGT
R F A V E N K T L S A L N I R E S
700
GACTTCTGGCAGCCTTTAACCCGTGCCGTGATGTTCAGCCAGCCTGCCAG
D F W Q P G T R A V M F S Q P A S
750
CCAGTTACTGGCAGGTGCGCGCATGGATGTGTATGTCATCCGTGACGGGG
Q L L A G A R M D V Y V I R D G
RBS
Stop Start(traB)
Primer 2
800
AGGGCAACtgatgGCCAGTATCAATACCATTGTGAAACGCAAGCAGTACC
E G N *
TGTGGCT

sequenced in order to guard against mutations introduced by *Taq* polymerase during PCR.

III. Measurement of termination activity within the *traK* gene

The highly polar nature of the *traK4* mutation suggested the presence of a rho-dependent terminator which would affect expression of downstream genes in the *tra* operon. Termination activity was measured using the vector pKL200 (McKenney *et al.*, 1981), which contains a multiple cloning site flanked by the *lac* promoter upstream and the *galK* gene downstream. Insertion of a fragment containing a transcription terminator results in a decrease in *galK* activity proportional to the efficiency of the terminator. The *traK* genes from pSPK1, pSPK4 and pSPK5 were inserted into pKL200 using *Bam*HI/*Eco*RI restriction sites and the level of galactokinase activity was measured in an *E. coli* HB101 background. The results of the *galK* assays are shown in Table 2.1. Measurements of *galK* activity in these constructs showed greatly reduced *galK* activity both for *traK4* and *traK105* compared with the wild-type gene. The low *galK* activity measured for *traK105* was unexpected, as this mutation has not been demonstrated to be polar in the F plasmid. It was therefore not predicted to cause termination of transcription. Interestingly, pKL200 containing the wild type *traK* gene showed strong *galK* activity, approximately 4-fold greater than the vector alone. To test for the presence of a promoter near the end of *traK*, we inserted the gene in the promoter assessment vector, pKO4 (McKenney *et al.*, 1981) and measured the resulting *galK* activity. No difference in *galK* activity could be detected between this construct and the vector alone, suggesting the absence of a promoter in *traK*.

IV. Computer-predicted transcription termination elements (TTEs) in *traK*

Using the features of the computer program described by Alifano *et al.* (1991), the entire transfer region was searched for possible C>G bubbles, also known as TTEs. TTEs were defined as regions of at least 78 nucleotides which contained a higher % of cytosine residues than guanosine residues. It is noteworthy that it is the ratio of C residues to G residues, rather than the absolute number of C's that is important. Sequential blocks of 78 nucleotides were searched at 1-bp intervals throughout the transfer region to identify sequences fitting these criteria. The search for TTEs in *traK* identified two such regions at nucleotides 131 to 330 and 429 to 510, the second of which lay between the *traK4* and *traK105* mutations (Figures 2.2 and 2.3). In addition, a number of these motifs occurred throughout the *tra* operon and are depicted in Figure 2.1. Strong TTE motifs (large arrows) were characterized by a percentage of C residues at least twice that of G residues

Table 2.1

A comparison of *galK* activity observed for wild type, K4 and K105 *tra*K genes assayed using a termination assessment vector

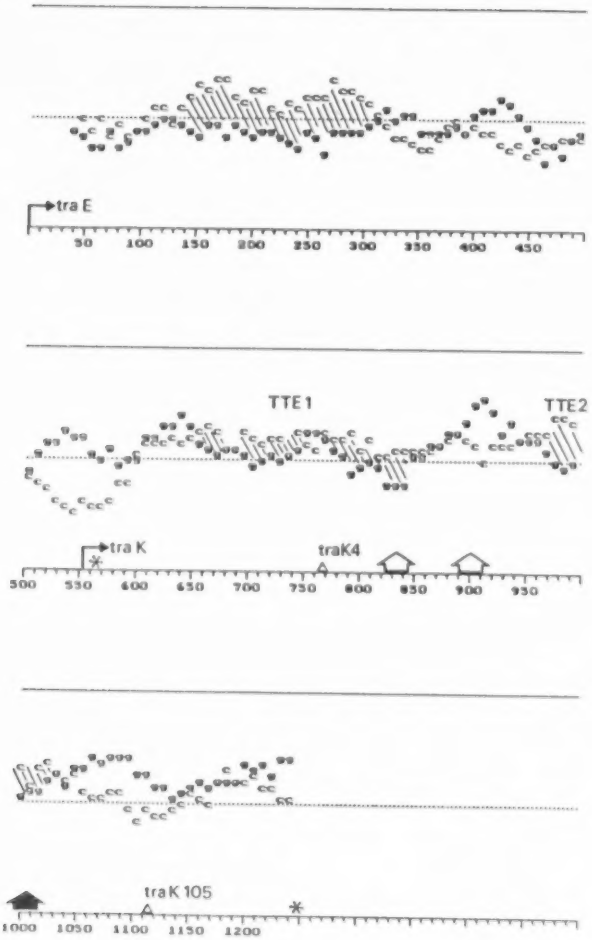
Construct	*Gal*K activity[a]	% *gal*K activity[b] compared with wild type
pKL200 (vector)	14	25
Wild Type	57	100
K4	5	9
K105	9	16

[a] The units were expressed as nanomoles of galactose phosphorylated per minute per unit of optical density of cells at 650nm.

[b] Three assays were performed and representative values from a single assay are given.

Figure 2.3

Computer analysis of the *traEK* genes of the F plasmid for potential transcription termination elements (TTEs). The dotted line represents 25% cytosine or guanosine content for each 78 nucleotide block searched. The *traK4* and *traK105* mutations are indicated (Δ) as well as the position of the 3' ends of protected fragments found by S1 analysis of wild type and K105 *traK* (open arrows) and *traK4* (filled arrow). Asterisks identify the translation stop points for the *traE* and *traK* genes. TTEs are shaded.


tra E
50
100
150
200
250
300
350
400
450
TTE1
TTE2
tra K
traK4
500
550
600
650
700
750
800
850
900
950
traK 105
1000
1050
1100
1150
1200

while weaker TTEs (small arrows) contained a greater percentage of C's than G's but the %C was less than two times greater than the %G. The two TTEs in the *traK* gene were considered to be weak signals by these criteria.

V. Analysis of the effectiveness of the TTEs in *traK*

The second TTE in *traK*, positioned between the *traK4* and *traK105* mutations, would be expected to truncate mRNA that is not actively translated. If this TTE is recognised under these conditions, a truncated transcript would be predicted for *traK4* but not *traK105* since translation continues downstream of the TTE in the latter mutant. Because of the extreme length of the *tra* operon transcript in the F plasmid itself, the fate of the RNA in the wild type and mutant plasmids could not be assessed easily. Therefore, the *traK* genes from the pSPK constructs were recloned into pTTQ118 and pTTQ119 such that they were under the control of an IPTG-inducible *tac* promoter in the vector, generating pQSK1, pQSK4 and pQSK5. S1 nuclease protection experiments were performed as described in Materials and Methods in order to define the 3' ends of the mRNA resulting from induction of the *tac* promoter in each construct. A protected fragment which terminated 290 bp from the *Bst*EII site was identified for pQSK4 (Figure 2.4); a small amount of full-length probe corresponding to the 337-bp *Bst*EII-*Pst*I fragment was detected upon over-exposure of the gel. Since an equal number of counts were loaded in each lane of Figure 2.4, these results indicate that most of the *traK4* transcripts terminated near the 3' end of TTE2. The protected fragments detected for mRNAs induced from pQSK1 and pQSK5 were identical; the strongest band represented the full length probe while weaker bands corresponding to processed transcripts ending 114 and 183 bp past the *Bst*EII site were also detected (Figure 2.4). Assuming that the transcripts began within the vector at the *tac* promoter, the ends of shorter transcripts detected in pQSK1 and pQSK5 samples lie between TTE1 and TTE2 while the 3' end of the truncated product in pQSK4 is located near the end of TTE2, suggesting efficient termination at this site.

The mRNA induced from the pQSK clones was subjected to Northern blot analysis in order to estimate its size and abundance (Figure 2.5). The pattern of degradation and/or processing of the mRNA in pQSK5 contained only minor differences from the pattern in pQSK1. The largest transcripts detected were approximately 880 nt long, corresponding to the predicted length for full-length RNA. A very weak signal corresponding to a transcript of approximately 540 nt long was detected for pQSK4 on overexposure of the autoradiograph. The size of this fragment is within the expected size range for a

Figure 2.4

traK4 transcripts are prematurely terminated. S1 nuclease analysis was performed on RNA extracted from cells expressing wild type *traK*, *traK4* and *traK105* genes. C indicates the sequence of C residues obtained from Maxam and Gilbert sequencing of the fragment used as a probe. FL indicates the position corresponding to the full length 337bp *Bst*EII / *Pst*I fragment (nt 196-533 in Figure 2.2) fragment used as a probe. The size of protected fragments is shown in base pairs. The autoradiogram is overexposed in order to visualize the protected fragments.

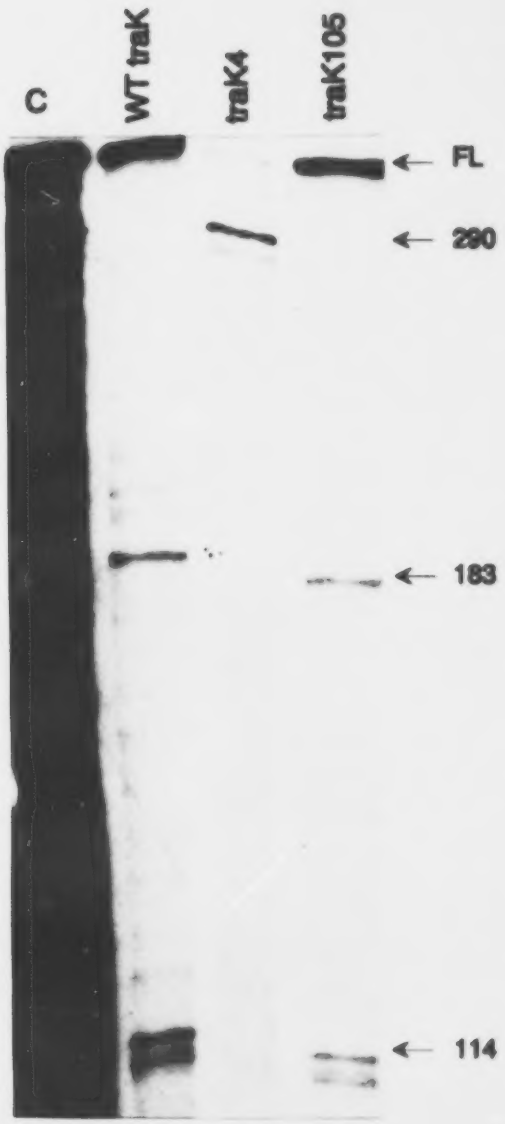
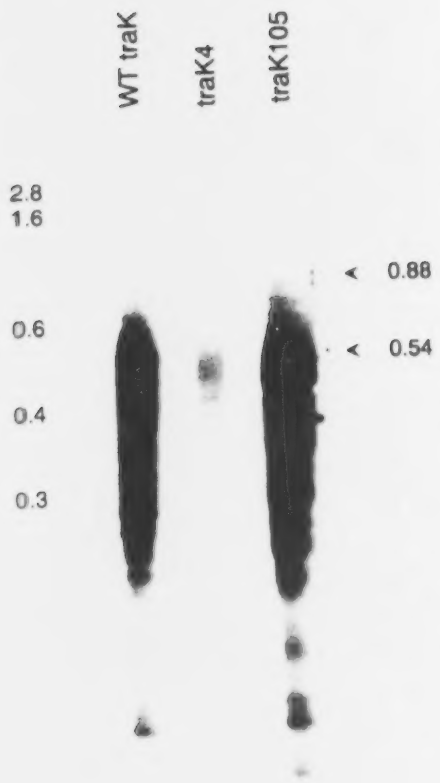

C
WT traK
traK4
traK105
FL
290
183
114

Figure 2.5

Northern blot analysis of the wild type *traK*, *traK4* and *traK105* transcripts. Radiolabelled oligonucleotide SPE4 was used to probe 10μg of RNA, separated on an 8% denaturing polyacrylamide gel. RNA size markers are in kilobases.


WT traK
traK4
traK105
2.8
1.6
0.6
0.4
0.3
0.88
0.54

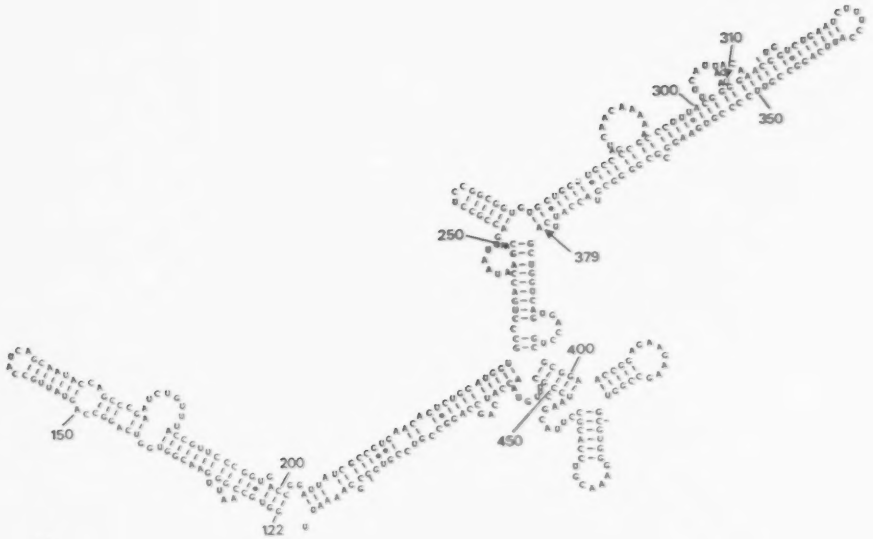
transcript expressed from the *tac* promoter and terminating at nt 486 (Figure 2.2) and thus was thought to correspond to the protected fragment detected by S1 analysis.

VI. *traK* mRNA secondary structure analysis

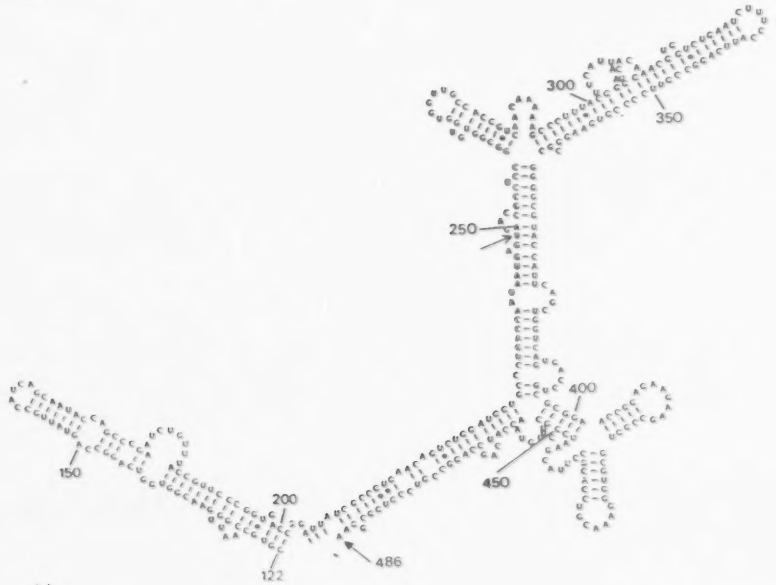
The overall increase in the rate of degradation of *traK4* RNA may be the result of changes in secondary structure. To investigate whether either the *traK4* or *traK105* mutations affected the secondary structure of the RNA, we made use of the computer algorithm RNAFOLD (Zuker and Stiegler, 1981) to predict the secondary structures of these transcripts. Nucleotides 122 to 490 of the wild type *traK* (and *traK105*, since the sequence is identical in this region) and nt 122 to 486 of the *traK4* sequences were folded into their most stable predicted configurations (Figure 2.6). The C-T transition at nt 249 in *traK4* was found to alter the predicted structure of the RNA between nt 241 and 381. Despite their different structures, both molecules were predicted to have similar calculated free energies of approximately -193 kcal (ca. -807.5kJ). The one base pair change that defines the *traK4* mutation may affect the pattern of degradation of the transcript by affecting the overall structure of the RNA. In contrast, the secondary structure predictions for nt 402 to 760 of the wild type and *traK105* transcripts, which encompass the *traK105* mutation, were identical.

Figure 2.6

The predicted secondary structure of the wild type *traK* / *traK105* (a) and *traK4* (b) transcripts from nucleotides 122-490 and 122 to 486 respectively, from Figure 2.2. The 3' ends of the protected fragments detected by S1 nuclease analysis are indicated by filled arrows (nt 310 and 379 in *a* and nt 486 in *b*) and the site of the *traK4* mutation is indicated by the open arrow. The secondary structure was generated by RNAFOLD (Zuker and Stiegler, 1981) using the PC/Gene program (Intelligenetics, Mountain View, California) and visualized by Loopviewer (Gilbert, 1990).


310
300
350
250
379
400
150
450
200
122

a)


300
350
250
400
150
450
200
486
122

b)

C. Discussion

Unlike rho-independent terminators which usually occur between genes, rho-dependent terminators have been identified both between genes, as in λ *tR1* (von Hippel, *et al.*, 1984) and *trpt'* (Platt, 1986), and within an operon, as in the *his* operon of *Salmonella typhimurium* (Alifano *et al.*, 1988; Ciampi *et al.*, 1989), and the *ilvGMEDA* operon (Wek *et al.*, 1987), and the *lacZ* gene of *E. coli* (Ruteshouser and Richardson, 1989). In the latter two cases, the polarity of certain mutations has been attributed to rho-dependent termination under conditions where translation has stopped. It has been proposed that the recognition of latent intragenic terminators under such conditions is a general mechanism to reduce unnecessary transcription (Richardson, 1990, 1991). Because of the large size of the *tra* operon, regulatory elements such as rho-dependent intragenic terminators would be expected to be present within this operon. The polar nature of the *traK4* mutation suggested that a rho-dependent terminator may have been unmasked and that this was the cause of the poor expression of genes downstream from *traK*. While *traB* is predicted to have a strong ribosome binding site, expression of this gene is greatly affected by the *traK4* mutation, suggesting that the primary effect of the mutation is at the transcriptional rather than the translational level.

With use of the PROSITE and SURFACEPLOT programs, TraK is predicted to have a molecular mass of 23,307 Da after processing from the original product of 25,627 Da. In previous analyses of *tra* products, only one polypeptide of approximately 24,000 Da has been identified and it was weakly expressed in all cases (Achtman *et al.*, 1979; Ippen-Ihler *et al.*, 1984; Thompson and Achtman, 1979); no precursor polypeptide was detected. Resolution of *tra* proteins in this size range is difficult, due to the large number of *tra* proteins close to the 25,000-Da range; this may explain the inability to detect a precursor product. Attempts to overexpress the TraK protein from the *traK*-containing fragment of pQSK1 placed under the control of an inducible T7 RNA polymerase promoter were unsuccessful. This suggests that TraK expression may be influenced by the presence of one or more *tra* genes within the larger fragments used previously (Achtman *et al.*, 1979; Ippen-Ihler *et al.*, 1984; Thompson and Achtman, 1979) to detect the *traK* gene product.

In galactokinase assays, reduced *galK* activity could be demonstrated with both the *traK4* and *traK105* genes. While reduced *galK* activity was expected for the *traK4* gene, the activity for *traK105* was unexpectedly low. This reduced activity appears to be the result

of reduced translation rather than transcriptional termination, since S1 nuclease analysis and Northern blot analysis demonstrated the presence of full-length *traK* transcripts in this construct. The *traK105* mutation could affect ribosome loading of the *galK* gene by allowing an alternative secondary structure to form in the absence of translation at the 3' end of this transcript. Also, the origin of the increased *galK* activity found for the wild type *traK* gene is unknown, since no promoter activity could be detected using the promoter assessment vector pKO4 (data not shown) and no fortuitous promoter was found at the cloning junction in pKL constructs.

The current model for the mechanism of rho-dependent termination suggests that these TTEs function by coupling transcription to translation to ensure that transcription does not continue when translation has been aborted (Richardson, 1991). In the absence of translation, specific sequences, called *rut* sequences, which are recognized by Rho protein, are exposed in the nascent RNA. Rho binds to the rut sequences and causes termination at a point just downstream of *rut* by a mechanism which is still unclear. In 1991, Alifano *et al.* (1991) described a motif which was common to all rho-dependent terminators and is proposed to constitute a *rut* site: a region of high cytosine over guanosine content called a C>G-rich bubble.

A search for TTE motifs in the *traK* gene revealed two of these transcription termination elements, one of which lies between the *traK4* and *traK105* mutations. The location of this TTE in *traK* supports the theory that the polar nature of the *traK4* mutation is due to rho-dependent termination. Abortion of translation at codon 73 in *traK4* would result in recognition by Rho of TTE2, and termination of transcription would occur. In *traK105*, none of the TTEs were predicted downstream of the translation stop codon suggesting that transcription would proceed into the next gene.

Nuclease protection experiments identified truncated transcripts in *traK4* which terminated at the 3' end of TTE2. In contrast, the majority of transcripts identified in wild type *traK* and *traK105* samples protected the probe from digestion by S1, suggesting that these transcripts were not terminated prematurely. The truncated transcripts detected for *traK4* in Northern blot analysis appeared unstable and were very rapidly degraded, possibly due to an altered secondary structure of the prematurely terminated transcripts which triggered ribonucleolytic degradation. In comparison, transcripts isolated from cells expressing cloned wild type *traK* and *traK105* were considerably more stable and fragments corresponding to full length transcripts were detected. It is possible that in the

F plasmid itself, alternate secondary structures and degradation patterns, differing from those described here, may occur when these genes are part of the large *tra* operon.

The RNA secondary structure predictions were identical for the wild type and *traK105* mRNA, but an alternate structure was predicted for *traK4* mRNA. It is interesting to note that a single base change in the *traK4* mutant significantly altered the predicted secondary structure. This altered structure could result in rapid degradation of the *traK4* transcripts by an alternate pathway to that used for wild type or *traK105* transcripts. One of the shorter transcripts observed in the wild type and *traK105* mutant (nt 379 in Figure 2.2) ends at the 3' side of a predicted stem-loop structure, suggesting that this may be a polymerase pause or RNase processing site. In *traK4*, this stem-loop no longer occurs, explaining the absence of an RNA transcript of corresponding size. These results suggest that subtle changes in sequence (1bp) can greatly affect RNA stability either by altering target sequences for ribonucleolytic cleavage or by decreasing chemical stability in the absence of translation.

The absence of any full length transcript from the *traK4* gene supports the theory that the polarity exhibited by this mutant is due to the recognition of a rho-dependent terminator in *traK* and computer analysis to detect possible terminators suggests that TTE2 could fulfill this role. Since the consensus sequence for *rut* sites can be characterized as having low stringency, this mechanism of termination could be widely used in a variety of genes. That a number of TTEs were found throughout the *tra* operon and were especially prominent in the longer genes such as *traE*, *-B*, *-C*, *-U*, *-N*, *-G*, and *-D* and *trbH* suggests that they play a role in terminating unnecessary transcription in this long operon. These results provide another example of how intragenic elements may function to control gene expression.

D. Materials and methods

I. Bacterial strains, plasmids and media.

Bacterial strains used in transformations were *E. coli* DH5α (*supE44* Δ*lacU169* (ϕ80 *lacZΔM15*) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*), MV1193 (Δ(*lac-proAB*) *rpsL thi endA sbcB15 hsdR4* Δ(*srl-recA)306*::Tn10(*tet*^{r}) F' [*traD36 proAB*$^+$ *lacI*^{q} *lacZΔM15*]) , BL21 (*hsdS gal* [λ*c*Its857 *ind*1 *S*am7 *nin5 lacUV5*-T7 gene1]) and HB101 (*supE44 hsdS20 recA13 ara-14 proA2 lacY1 galK2 rpsL20 xyl-5 mtl-1*). M176, used as the donor strain for mating assays, contains the F plasmid, JCFL0 (Achtman *et al.*, 1971), in *E. coli* JC3272 (F$^-$ *lac* ΔX74 *galK his trp lys str* λ^r T6^r), while *E. coli* ED24 (F$^-$ Lac$^-$ Spc^R) was the recipient. Plasmids F*lac traK4* (JCFL4) and *Flac traK105* (JCFL105) have been described (Achtman *et al.*, 1971; Willetts and Achtman, 1972) and were supplied by N.S. Willetts and K. Ippen-Ihler respectively, while pKO4 and pKL200 were obtained from N. Willetts and are related to plasmids developed by McKenney *et al.* (1981). Plasmids pTTQ18 and 19 (Stark, 1987) were purchased from Amersham and LB (Luria-Bertani) broth was as previously described (Maniatis *et al.*, 1982).

II. Recombinant DNA techniques and reagents

Restriction and DNA modification enzymes were purchased from Boehringer-Mannheim and, unless otherwise stated, used according to the manufacturer's instructions. The wild type *traK* gene was subcloned from pRS27 (Skurray *et al.*, 1978) using convenient restriction sites (Frost *et al.*, 1994) into M13mp18 (Yanisch-Perron *et al.*, 1985) by standard methods and sequenced by the chain termination method (Sanger *et al.*, 1977) using a Sequenase sequencing kit (United States Biochemical Corporation). Two primers of sequence 5'-GACCTGGCTGGATAATTTCG-3' and 5'-AGCCACAGGTACTGCTTGCG-3' (Figure 2.2) were used in the polymerase chain reaction to amplify an 850 bp fragment containing the entire *traK* gene from purified F plasmid DNA isolated from M176 and JC3272 containing each of the two mutant F plasmids. The amplified fragments were then treated with Klenow and T4 polynucleotide kinase and were ligated into *Sma*I-digested pUC118 (Vieira and Messing 1987). These constructs were transformed into *E. coli* MV1193 and single-stranded DNA was prepared for sequencing (Vieira and Messing, 1987). Plasmids pSPK1 contained the wild type *traK* gene oriented away from the *lac* promoter, while pSPK4 and pSPK5 contained the *traK4* and *traK105* genes oriented in the same direction as the *lac* promoter, respectively. The *Bam*HI/*Eco*RI fragments from pSPK1, pSPK4 and pSPK5 were cloned into pTTQ18

and 19 vectors (Stark, 1987) to create plasmids pQSK1, pQSK4 and pQSK5, respectively, such that the *traK* gene was aligned with the *tac* promoter.

III. Assay for transfer efficiency

Plasmid transfer ability was quantitated using the procedure described in Frost *et al.*(1989). Donor (*E. coli* JC3272 containing wild type JCFL0, *Flac traK105* or JCFL*traK4* and one of pSPK1, pSPK4 or pSPK5) and recipient cells (*E. coli* ED24) were grown to mid-log phase in LB broth at 37°C and 0.1 ml of each was mixed together with 1.0 ml of fresh broth. The cells were incubated at 37°C for 1 hour, diluted 100-fold in 1xSSC, vortexed to interrupt mating and plated on selective media after further dilution.

IV. *galK* assay

The assay to measure *galK* activity was performed as described (McKenney *et al.*, 1981) except that the specific activity of the [^{14}C]galactose was 59.6 mCi/mmol and the background strain was *E. coli* HB101. The units were expressed as nanomoles of galactose phosphorylated per minute per unit of optical density of cells at 650 nm.

V. Computer-predicted termination sites

Using the parameters suggested by Alifano *et al.* (1991), a computer program for the IBM PC was written in Turbo Pascal by Ken Usher, Department of Microbiology, University of Alberta to assess the presence of possible termination sites within the entire F transfer region. Transcription Termination Elements (TTEs) were defined as regions of at least 78 nucleotides where the % of cytosine residues was greater than the % of guanosine residues.

VI. RNA preparation

Cells (1.5 ml) containing pQSK1, pQSK4 or pQSK5 were induced for three minutes by the addition of IPTG (isopropylthiogalactoside) to a final concentration of 0.5 mM followed by the addition of rifampicin to a final concentration 200μg.ml^{-1} for 2 minutes. RNA was prepared by the modified hot phenol extraction method described previously (Frost *et al.*, 1989).

VII. S1 nuclease analysis

Nuclease protection experiments were carried out using the 337 bp *Bst*EII/*Pst*I (Figure 2.2) fragment as a probe. The fragment was labelled with [α^{32}P]-dATP using Klenow

enzyme to "fill in" the recessed 3' end of the *Bst*EII site (Maniatis *et al.*, 1982) and purified by passage through a Biogel P-30 column. Hybridization of 0.1 pmol of probe to 50 µg total RNA was allowed to proceed overnight at 50°C prior to treatment with 400 units of S1 nuclease. Maxam and Gilbert sequencing reactions (Maxam and Gilbert, 1980) were carried out simultaneously on the labelled *Bst*EII/*Pst*I fragment and the reactions were electrophoresed on a 6% denaturing polyacrylamide gel.

VIII. Northern blot analysis

Oligonucleotide SPE4 (5'-GGCTGATGGTTACAAGCAGGG-3') is complementary to the 3' end of TTE2 (Figure 2.2) and was end-labelled with [γ^{32}P]-ATP using T4 polynucleotide kinase (Sambrook *et al.*, 1989). RNA extracted from *E. coli* containing pQSK1, pQSK4 and pQSK5 was subsequently probed with the labeled primer. RNA was separated by electrophoresis on a 5% polyacrylamide gel containing 8.5 M urea, and transferred to a Hybond N nylon membrane (Amersham) using a Biorad Trans-Blot Cell. The membrane was pre-hybridized for a minimum of 2 hours at 37°C in 2.5xSSC, 5x Denhardt's solution (Sambrook *et al.*, 1989), 1.5% SDS, 100 µg.ml^{-1} of *E.coli* strain W tRNA type XX (Sigma) and 100µg.ml^{-1} of denatured calf thymus DNA. Hybridization was at 56°C overnight and the membrane was washed with 6xSSC, 0.1% SDS for 2x10 minutes at 56°C and 2x10 minutes at 61°C. Autoradiography was performed at -70°C with an intensifying screen using Kodak X-AR5 film.

IX. Secondary structure analysis

Sequence analysis and secondary structure predictions were generated by RNAFOLD (Zuker and Stiegler, 1981) using the PC/Gene program (Intelligenetics, Mountain View, California) and the LoopViewer RNA secondary structure viewing program (Gilbert, 1990).

X. Genbank accession number

The nucleotide sequence of *traK* can be accessed through Genbank accession number U01159.

E. References

Achtman, M., P. A. Manning, C. Edelbluth, and P. Herrlich. (1979) Export without proteolytic processing of inner and outer membrane proteins encoded by F sex factor *tra* cistrons in *Escherichia coli* minicells. *Proc. Natl. Acad. Sci. USA* **76**: 4837-4841.

Achtman, M., N. S. Willetts, and A. J. Clark. (1971) Beginning a genetic analysis of transfer determined by the F factor in *Escherichia coli* by isolation and characterization of transfer-deficient mutants. *J. Bacteriol.* **106**: 529-538.

Achtman, M., N. Willetts, and A. J. Clark. (1972) Conjugational complementation analysis of transfer-deficient mutants of F *lac* in *Escherichia coli*. *J. Bacteriol.* **110**: 831-842.

Alifano, P., M. S. Ciampi, A. G. Nappo, C. B. Bruni, and M. S. Carlomagno. (1988) In vivo analysis of the mechanisms responsible for strong transcriptional polarity in a "sense" mutant within an intercistronic region. *Cell* **55**: 351-360.

Alifano, P., F. Rivellini, D. Limauro, C. B. Bruni, and M. S. Carlomagno. (1991) A consensus motif common to all Rho-dependent prokaryotic transcription terminators. *Cell* **64**: 553-563.

Ciampi, M. S., P. Alifano, A. G. Nappo, C. B. Bruni, and M. S. Carlomagno. (1989) Features of the Rho-dependent transcription terminator polar element within the *hisG* cistron of *Salmonella typhimurium*. *J. Bacteriol.* **171**: 4472-4478.

Frost, L. S., K. Ippen-Ihler, and R. A. Skurray. (1994) An analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol. Rev.* **58**: 162-210.

Frost, L., S. Lee, N. Yanchar, and W. Paranchych. (1989) *finP* and *fisO* mutations in FinP antisense-RNA suggest a model for FinOP action in the repression of bacterial conjugation by the F*lac* plasmid JCFLO. *Mol. Gen. Genet.* **218**: 152-160.

Frost, L. S., W. Paranchych, and N. S. Willetts. (1984) DNA Sequence of the F *traALE* region that includes the gene for F pilin. *J. Bacteriol.* **160**: 395-401.

Gilbert, D. G. (1990) *Loopviewer*, a Macintosh program for visualizing RNA secondary structure. Published electronically on the Internet, available via anonymous ftp to iubio.bio.indiana.edu.

Ham, L. M., D. Cram, and R. Skurray. (1989.) Transcriptional analysis of the F plasmid surface exclusion region: mapping of *traS*, *traT* and *traD* transcripts. *Plasmid* **21**: 1-8.

Ippen-Ihler, K., M. Achtman, and N. Willetts. (1972) Deletion map of the *Escherichia coli* K12 sex factor F: the order of eleven transfer cistrons. *J. Bacteriol.* **110**: 857-863.

Ippen-Ihler, K., and S. Maneewannakul. (1991) Conjugation among enteric bacteria: mating systems dependent on expression of pili, p. 35-69. *In* M. Dworkin (ed.), Microbial Cell-Cell Interactions. American Society for Microbiology, Washington D.C.

Ippen-Ihler, K., D. Moore, S. Laine, D. A. Johnson, and N. S. Willetts. (1984) Synthesis of F-pilin polypeptide in the absence of F *traJ* product. *Plasmid* **11**: 116-129.

Maniatis, T., E. F. Fritsch, and J. Sambrook. (1982) Molecular cloning, a Laboratory Manual, p113-114. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Maxam, A. M., and W. Gilbert. (1980) Sequencing end-labelled DNA with base specific chemical cleavages. *Meth. Enzymol.* **65**: 499-580.

McKenney, K., H. Shimatake, D. Court, U. Schmeissner, C. Brady, and M. Rosenberg. (1981) A system to study promoter and terminator signals recognized by *Escherichia coli* RNA polymerase, p. 383-415. *In* Chirikjian, J. and T. Papas (eds.), Gene Amplification and Analysis, Vol II. Elsevier-North Holland, New York.

Platt, T. (1986) Transcription termination and the regulation of gene expression. *Annu. Rev. Biochem.* **55**: 339-372.

Richardson, J. P. (1990) Rho-dependent transcription termination. *Bioch. Biophys. Acta.* **1048**: 127-138.

Richardson, J. P. (1991) Preventing the synthesis of unused transcripts by Rho factor. *Cell* **64**: 1047-1049.

Rivellini, F., P. Alifano, C. Piscitelli, V. Blasi, C. Bruni, and M. S. Carlomagno. (1991) A cytosine- over guanosine-rich sequence in RNA activates rho-dependent transcription termination. *Mol. Microbiol.* **5**: 3049-3054.

Ruteshouser, E. C., and J. P. Richardson. (1989) Identification and characterization of transcription termination sites in the *Escherichia coli lacZ* gene. *J. Mol. Biol.* **208**: 23-43.

Sambrook, J., E. F. Fritsch, and T. Maniatis. (1989) Molecular cloning. A laboratory manual (2nd ed), p. 5.68-5.69. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sanger, F., S. Nicklen, and A. R. Coulson. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**: 5463-5467.

Skurray, R. A., H. Nagaishi, and A. J. Clark. (1978) Construction and *Bam*HI analysis of chimeric plasmids containing *Eco*RI DNA fragments of the F sex factor. *Plasmid* **1**: 174-186.

Stark, M. J. R. (1987) Multicopy expression vectors carrying the *lac* repressor gene for regulated high level expression of genes in *Escherichia coli*. *Gene* **51**: 255-267.

Studier, F. W., and B. A. Moffat. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**: 113-130.

Tabor, S., and C. C. Richardson. (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* **82**: 1074-1078.

Thompson, R., and M. Achtman. (1979) The control region of the F sex factor DNA transfer cistrons: physical mapping by deletion analysis. *Mol. Gen. Genet.* **169**: 49-57.

Vieira J., and J. Messing. (1987) Production of single-stranded plasmid DNA. *Meth. Enzymol.* **153**: 3-11.

von Hippel, P. H., D. G. Bear, W. D. Morgan, and J. A. McSwiggen. (1984) Protein-nucleic acid interactions in transcription: a molecular analysis. *Annu. Rev. Biochem.* **53**: 389-446.

Wek, R. C., J. H. Sameshima, and G. W. Hatfield. (1987) Rho-dependent transcriptional polarity in the *ilvGMEDA* operon of wild-type *Escherichia coli* K12. *J. Biol. Chem.* **262**: 15256-15261.

Willetts, N. (1977) The transcriptional control of fertility in F-like plasmids. *J.Mol. Biol.* **112**: 141-148.

Willetts, N. and M. Achtman. (1972) Genetic analysis of transfer by the *Escherichia coli* sex factor F, using P1 transductional complementation. *J. Bacteriol.* **110**: 843-851.

Willetts, N., J. Maule, and S. McIntire. (1976) The genetic locations of *traO*, *finP* and *tra-4* on the *E.coli* sex factor F. *Genet. Res. Camb.* **26**: 255-263.

Willetts, N., and R. Skurray. (1987) Structure and function of the F factor and mechanism of conjugation, p. 1110-1133. *In* F.C. Neidhart, J.L. Ingraham, K.B. Low, B. Magasanik, M. Scheacter, and H.E. Umbarger, (eds.), *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology. American Society for Microbiology, Washington.

Yager, T. D., and P. H. von Hippel. (1991) A thermodynamic analysis of RNA transcript elongation and termination in *Escherichia coli*. *Biochemistry* **30**:1097-1118.

Yanisch-Perron, C., J. Vieira, and J. Messing. (1985) Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-109.

Zuker, M., and P. Stiegler. (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxilliary information. *Nucleic Acids Res.* **9**: 133-148.

Chapter 3

An analysis of the expression of the *traM* gene of the F sex factor of *Escherichia coli*

A. Introduction

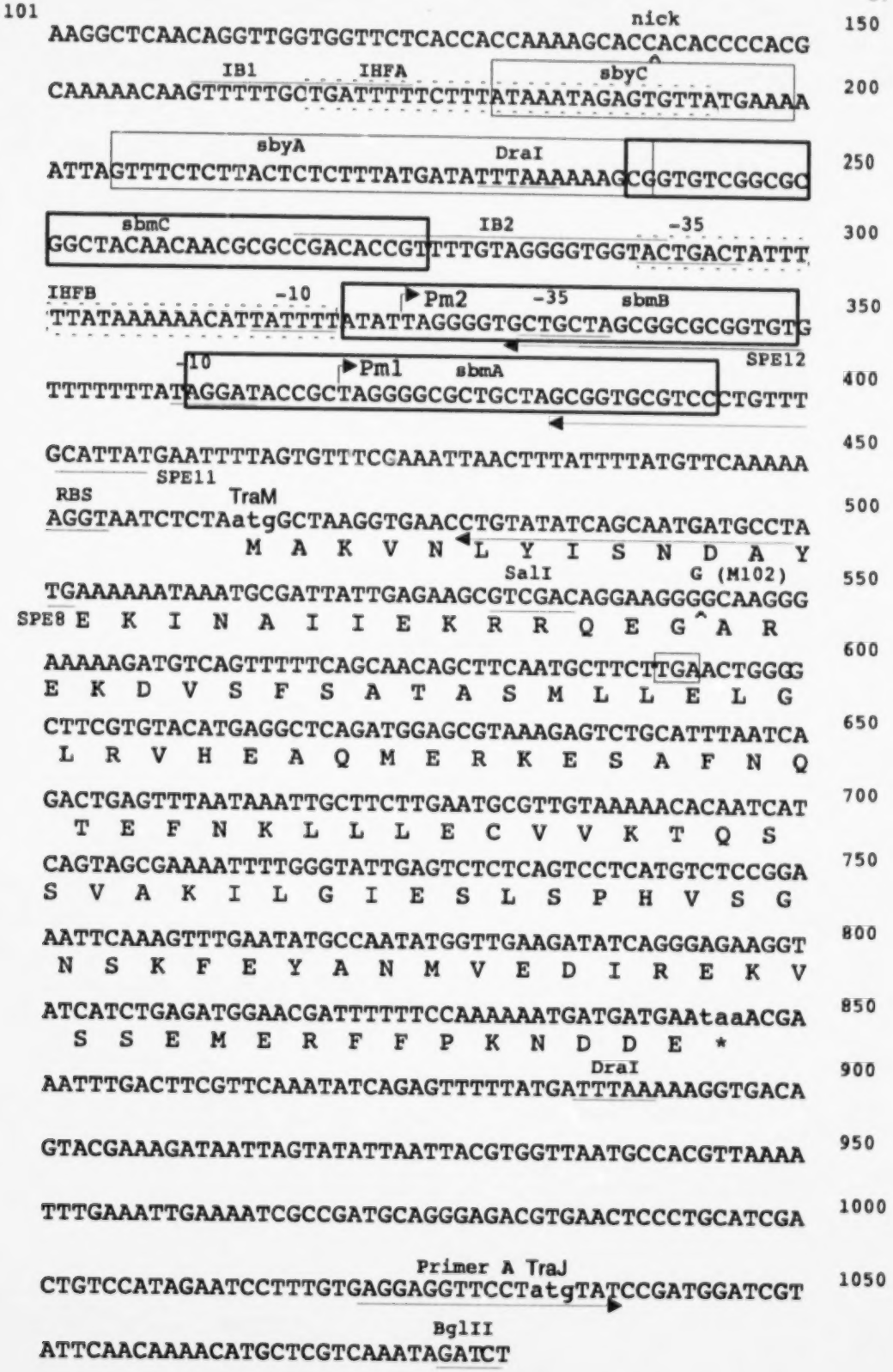
The regulation of bacterial conjugation appears to be a complex system, involving the control of expression of the 28 genes known to be required for this process (Frost *et al.*, 1994; Ippen-Ihler and Maneewannakul, 1991; Willetts and Skurray, 1987). Expression of the *tra* operon is positively controlled by the *traJ* gene product which in turn is negatively controlled by the *finOP* antisense RNA system. Outside of the *tra* operon, lies the origin of transfer, where strand- and -site specific nicking occurs and strand transfer is initiated, and *traM*, the gene product of which has tentatively been assigned the role of signalling that a stable mating pair has been formed and that transfer can begin (Kingsman and Willetts, 1978; Willetts and Wilkins, 1984). TraM has been shown to be a DNA-binding protein which binds to three sites within the *oriT* of F (Figure 3.1) (Di Laurenzio *et al.*, 1992). The highest affinity binding site (*sbmA*), as determined by DNaseI footprinting, overlaps the predicted start of *traM* transcription, while the second highest affinity binding site, *sbmB*, lies within the -35 region. The site with the lowest affinity for TraM, *sbmC*, lies between the TraY binding site within *oriT* (Lahue and Matson, 1990) and an integration host factor (IHF) binding site, IHFB (Tsai *et al.*, 1990).

The regulation of transcription of *traM* is not well understood. Conflicting reports exist on the effect of *traJ* on *traM* transcription in the F plasmid. Using *lacZ -traM* fusions, Gaffney *et al.* (1983) showed that *traJ* was required for the expression of *traM*, while Mullineaux and Willetts (1985) have reported that *traJ* had no effect on the transcription of *traM* when the *traM* promoter was fused to *galK*. However, all plasmids expressing *traJ* also expressed *traM*, which makes interpretation of this work difficult, since the autoregulatory nature of *traM* was not considered.

The antibiotic resistance plasmid R100-1 has been shown to contain two overlapping *traM* transcripts (Dempsey, 1989). The longer of these two transcripts was maximally expressed only when present on the derepessed R-100-1 plasmid, implying that either gene products encoded by this plasmid, or a particular DNA structure determined by this plasmid is required for efficient *traM* expression. More recently, a third, constitutively expressed promoter has been identified in R100 (Abo and Ohtsubo, 1993), and the major *traM* promoter was shown to be repressed by the TraM protein. In addition, this promoter was not affected by the presence of a co-resident plasmid expressing *traJ*, thus agreeing with the results of Mullineaux and Willetts (1985). Integration host factor (IHF) is a chromosomally encoded DNA-binding protein of *E. coli*, composed of two non-

Figure 3.1

The nucleotide sequence of the 1.1 kb *Bgl*II fragment of F, containing the nick site, *oriT*, and the *traM* gene. The three binding sites previously identified for the TraM protein, *sbmA*, *sbmB*, and *sbmC* are boxed, as are the two IHF binding sites (IHFA and IHFB) and the TraY binding sites (*sbyA* and *sbyC*). The 5' ends of transcripts identified in this work are indicated (Pm1 and Pm2), together with proposed -10 and -35 sequences, which are underlined. IB1 and IB2 are sequence directed bends in *oriT*. The ribosome binding site (RBS) for *traM* is underlined. Lower case letters represent the translation initiation codons of *traM* and *traJ*, and the translation termination codon of *traM*. The frameshift mutation identified for M102 is indicated at nt 541 and the termination codon that occurs as a result of this frameshift mutation is boxed. Primers discussed in this chapter are indicated and the amino acid sequence of *traM* is given below the nucleotide sequence.


101
nick
150
AAGGCTCAACAGGTTGGTGGTTCTCACCACCAAAAGCACCACACCCCACG
IB1
IHFA
sbyC
200
CAAAAACAAGTTTTTGCTGATTTTTCTTTATAAATAGAGTGTTATGAAAA
sbyA
DraI
250
ATTAGTTTCTCTTACTCTCTTTATGATATTTAAAAAAGCGGTGTCGGCGC
sbmC
IB2
-35
300
GGCTACAACAACGCGCCGACACCGTTTTGTAGGGGTGGTACTGACTATTT
IHFB
-10
Pm2
-35
sbmB
350
TTATAAAAAACATTATTTTATATTAGGGGTGCTGCTAGCGGCGCGGTGTG
SPE12
-10
Pm1
sbmA
400
TTTTTTTATAGGATACCGCTAGGGGCGCTGCTAGCGGTGCGTCCCTGTTT
450
GCATTATGAATTTTAGTGTTTCCAAATTAACTTTATTTTATGTTCAAAAA
SPE11
RBS
TraM
500
AGGTAATCTCTAatgGCTAAGGTGAACCTGTATATCAGCAATGATGCCTA
M A K V N L Y I S N D A Y
SalI
G (M102)
550
TGAAAAAATAAATGCGATTATTGAGAAGCGTCGACAGGAAGGGGCAAGGG
SPE8 E K I N A I I E K R R Q E G A R
600
AAAAAGATGTCAGTTTTTCAGCAACAGCTTCAATGCTTCTTGAACTGGGG
E K D V S F S A T A S M L L E L G
650
CTTCGTGTACATGAGGCTCAGATGGAGCGTAAAGAGTCTGCATTTAATCA
L R V H E A Q M E R K E S A F N Q
700
GACTGAGTTTAATAAATTGCTTCTTGAATGCGTTGTAAAAACACAATCAT
T E F N K L L L E C V V K T Q S
750
CAGTAGCGAAAATTTTGGGTATTGAGTCTCTCAGTCCTCATGTCTCCGGA
S V A K I L G I E S L S P H V S G
800
AATTCAAAGTTTGAATATGCCAATATGGTTGAAGATATCAGGGAGAAGGT
N S K F E Y A N M V E D I R E K V
850
ATCATCTGAGATGGAACGATTTTTTCCAAAAAATGATGATGAATaaACGA
S S E M E R F F P K N D D E *
DraI
900
AATTTGACTTCGTTCAAATATCAGAGTTTTTATGATTTAAAAAGGTGACA
950
GTACGAAAGATAATTAGTATATTAATTACGTGGTTAATGCCACGTTAAAA
1000
TTTGAAATTGAAAATCGCCGATGCAGGGAGACGTGAACTCCCTGCATCGA
Primer A TraJ
1050
CTGTCCATAGAATCCTTTGTGAGGAGGTTCCTatgTATCCGATGGATCGT
BglII
ATTCAACAAAACATGCTCGTCAAATAGATCT

identical subunits encoded by the *himA* and *hip* genes. The protein has been demonstrated to be involved in a number of cellular processes, including recombination events, DNA replication and regulation of gene expression (reviewed in Friedman, 1988). Conflicting reports exist on the effect IHF on R100 *traM* transcription - in one case it was shown to represss *traM* (Abo and Ohtsubo, 1993), in the other it was shown to enhance *traM* transcription (Dempsey and Fee, 1990). In plasmid R1, two *traM* transcripts were identified by S1 nuclease analysis, with the longer of the two transcripts predominating (Koronakis *et al.*, 1985). It has been shown that *traM* of this plasmid is autoregulated (Schwab *et al.*, 1993) but no effect of *traJ* on *traM* expression could be detected.

In vitro transcription experiments (Thompson and Taylor, 1982) have shown the promoter for F *traM* to lie 165 bp upstream from the *Sal*I site within the *traM* gene (Figure 3.1). The presence of more than one promoter in the *traM* genes of plasmids R1 and R100 suggested to us that a detailed transcriptional analysis of *traM* expression in the F plasmid might reveal the presence of additional promoters. This chapter describes an analysis of the transcriptional regulation of F *traM* to determine which *tra*- and host-encoded genes affect the expression of this gene. Two *traM* promoters were identified, one of which corresponds to the promoter identified by Thompson and Taylor. In addition, we confirm the finding of others (Schwab *et al.*, 1993; Abo and Ohtsubo, 1993) that *traM* expression is autoregulated. We also investigated the expression of TraM protein from a variety of *tra* mutants and found that the results correlated well those obtained in the transcriptional analysis.

B. Results

I. Construction of pOX*traM*K3

The *traM* gene in plasmid pLDLF7 (Di Laurenzio *et al.*, 1992) was interrupted by the insertion of a kanamycin resistance cassette at the *Sal*I site, to create pMF7-Kan. The mutant *traM* gene was crossed into pOX38, a transfer proficient derivative of F, in a triparental mating experiment (Figure 3.2). Analysis of the plasmid content in a sample of the apparent transconjugants indicated the presence of pMF7-Kan DNA, suggesting that co-integrates had formed, thus allowing the transfer of this non-conjugative plasmid. Mobilization of non-conjugative plasmids by pOX38, which lacks all known insertion sequences and transposable elements except for a 200bp portion of IS3, has been described previously (O'Connor and Malamy, 1984). Replica plating onto ampicillin-containing plates identified 10/360 transconjugants that were sensitive to ampicillin. These transconjugants were also shown to be sensitive to f1 phage, indicating that they expressed functional F pili. Small-scale DNA analysis (Birnboim and Doly, 1979) showed that these cells did not contain pMF7-Kan DNA, so these were analysed further to confirm the presence of the kanamycin resistance cassette within the *traM* gene of pOX38. Plasmid DNA extracted from one of these transconjugants and digested with *Bgl*II was subjected to hybridization analysis (Sambrook *et al.*, 1989) using end-labelled Primer A (Figure 3.1). An increase in the size of this fragment from 1.1 kb to 2.4 kb was observed, corresponding to an insertion of 1.3 kb (data not shown). This plasmid was called pOX*traM*K3.

II. Mating efficiencies of plasmids used in this study

The efficiency of transfer of the plasmids used in this study was compared with that of pOX38::Km (Chandler and Galas, 1983), which is identical to pOX38 except for the addition of a kanamycin resistance cassette at the *Hin*dIII site (Table 3.1). The requirement for IHF in transfer is demonstrated by a 350-fold decrease in transfer efficiencies of pOX38::Km in a *himA*$^-$, *hip*$^-$ host, MC253. A *traY* insertion mutant, pOX*traY*244, was transferred with an efficiency 2000-fold lower than the wild type plasmid. This mutation appeared not to be completely polar, as mating efficiencies of this plasmid could be increased 10-fold by suppyling TraY *in trans* from pRS27. The *traM* insertion mutant created in this work, pOX*traM*K3, was unable to support transfer. However, this mutant could be complemented by *traM* expressed in *trans* from a variety of multicopy plasmids. Transfer levels could be restored to half that of wild type levels when *traM* was expressed in *trans* from pLDLF7. pLFR28, which lacks any

Figure 3.2

Construction of pOX*traM*K3. The *DraI* fragment of pLDLF7 contains the entire *traM* gene, including the *traM* promoter (open box). The kanamycin resistance cassette (shaded box) was inserted within the *traM* gene of pLDLF7 to create pMF7-Kan and crossed into pOX38, which has a wild type *traM* gene. pOX*traM*K3 was then mated into *E. coli* XK1200 (naladixic acid resistant) and colonies resistant to both naladixic acid and kanamycin were selected for further analysis.

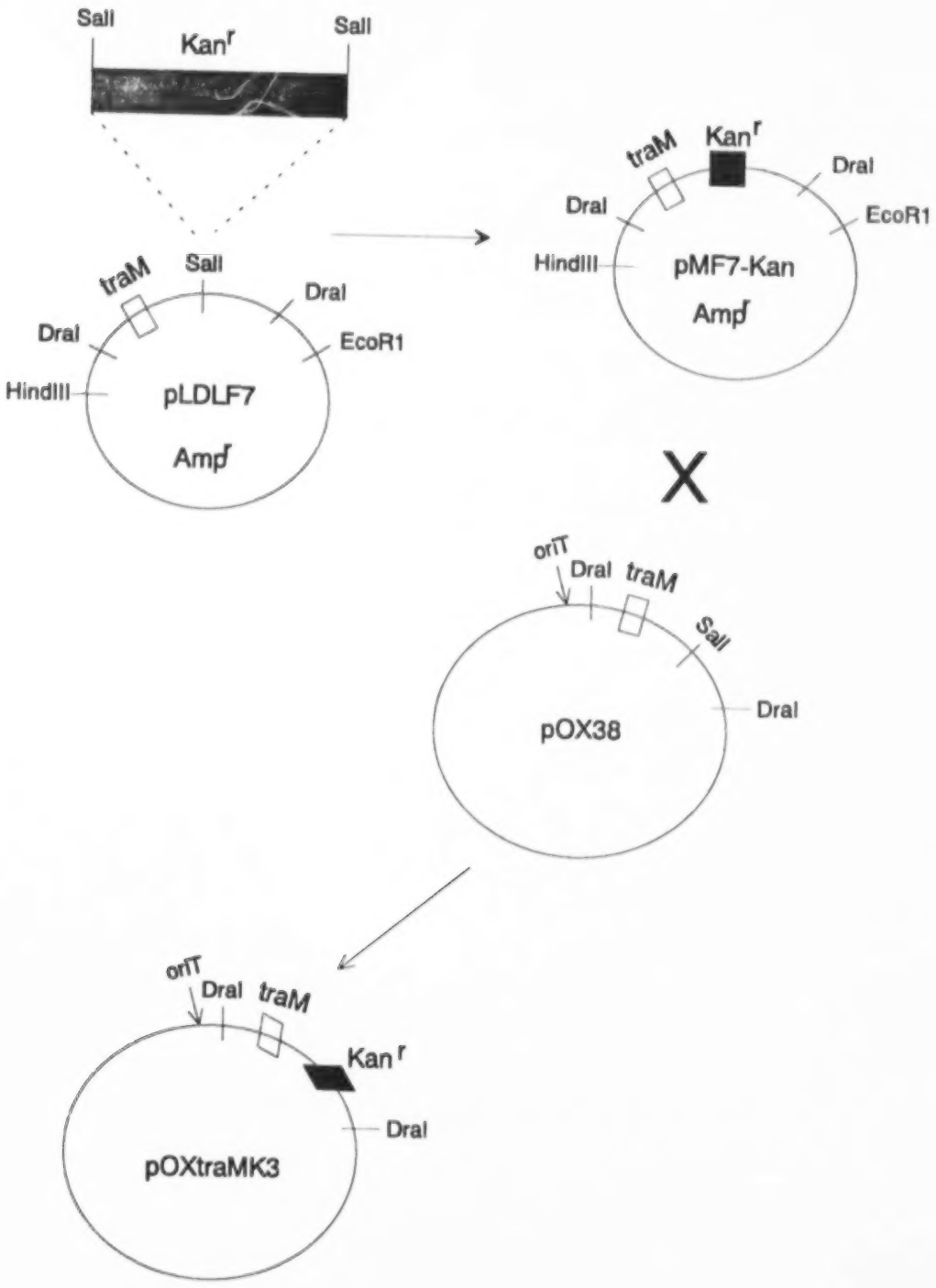

SalI
Kan^r
SalI
traM
SalI
DraI
DraI
EcoR1
HindIII
pLDLF7
Amp^r
traM
Kan^r
DraI
DraI
EcoR1
HindIII
pMF7-Kan
Amp^r
X
oriT
DraI
traM
SalI
DraI
pOX38
oriT
DraI
traM
Kan^r
DraI
pOXtraMK3

Table 3.1

Mating efficiencies of plasmid pOX38::Km and its derivatives.

Plasmid	Number of transconjugants/100 donors
pOX38::Km	22
pOX38::Km (MC253)	0.06
pOX*traY*244	0.01
pOXtraY244 + pRS27	0.1
pOX*traM*K3	0
pOX*traM*K3 + pNY300	40
pOX*traM*K3 + pLDLF7	11
pOX*traM*K3 + pLFR28	0.1

detectable *traM* promoters but expresses small amounts of TraM from a vector promoter (detectable by immunoblot analysis), also supported transfer when expressed in *trans*.

III. Mapping of the 5' ends of *traM* transcripts by primer extension analysis

Analysis of the 5' ends of *traM* transcripts was done using primer SPE8 which binds 118 nucleotides downstream of the predicted initiation site. For pOX38::Km, two transcripts were detected (Figure 3.3). The shortest transcript, (Pm1), initiated 370 nt downstream of the *Bgl*II site on the left of *oriT*. This corresponds to the initiation site that was identified by Thompson and Taylor (1982) for F *traM* by run-off transcription experiments using linear templates. The second *traM* transcript appeared to be a triplet, but prior treatment of RNA with DNaseI caused the disappearance of the longest of these cDNAs. In contrast, only treatment with RNase resulted in the disappearance of the lower two bands, confirming that these cDNAs reflected *traM* transcripts. The longest of these two RNAs initiated 324 nucleotides downstream of the *Bgl*II site (Pm2). The second transcript was likely a degradation product of the first, since its appearance and intensity was always correlated with that of the longer transcript. Both transcripts could be detected in all pOX38 plasmids tested, but the relative amounts of each transcript varied in the pOX38 mutants tested. In wild type pOX38::Km, the shortest transcript, initiating at Pm1, was more abundant than Pm2. Similarly, in pOX*traM*K3, transcripts initiating at Pm1 appeared to be more abundant. Densitometric analysis of the results from three separate experiments indicated that the intensity of bands representing Pm1 and Pm2 were at least five-fold higher in this construct than in pOX38::Km. A five-fold decrease in expression from both Pm1 and Pm2 was detected in pOX*traY*244 as compared to pOX38::Km. However, *traM* expression from Pm1 and Pm2 in a *traJ* mutant, F*lac traJ*90, decreased only two-fold in comparison to that in pOX38::Km, which shows comparable expression to the F plasmid from these two promoters. Since IHF has been shown to severely reduce mating efficiencies (Table 3.1), and is known to bind to the *oriT* region of F (Tsai *et al.*, 1990), we investigated whether IHF has an effect on *traM* transcription. No difference in the expression of *traM* from either Pm1 or Pm2 could be detected in the IHF-deficient strain, *E. coli* MC253.

Primer SPE8 was also used to analyse *traM* transcripts from multicopy plasmids containing cloned fragments of the *oriT* region (Figure 3.3). When RNA isolated from pNY300 (Frost *et al.*, 1989), a multicopy plasmid carrying the 1.1kb *Bgl*II insert shown in Figure. 3.4 was used as a template for primer extension, the same two *traM* promoters

Figure 3.3

Primer extension analysis to detect the 5' ends of F *traM* transcripts. RNA was extracted from *E. coli* XK1200 cells containing the indicated plasmids. G, A, T, C represent sequence derived from dideoxy sequencing reactions using pNY300 as a template and primer SPE8. F*lac traJ*90 shows the results of primer extensions done on RNA extracted from JC3272 cells containing the F *traJ* mutant, F*lac traJ*90. XK1200 is a control cell line containing no plasmid. Arrows represent the 5' ends of *traM* transcripts Pm1 and Pm2, mapped in this work. The right half of the figure represents the results of primer extension analysis from multicopy plasmids containing cloned regions of *oriT* and *traM*. pLFR28 lacks both *traM* promoters. Exposure times were one week for products from single copy plasmids, and overnight for results from multicopy plasmids.

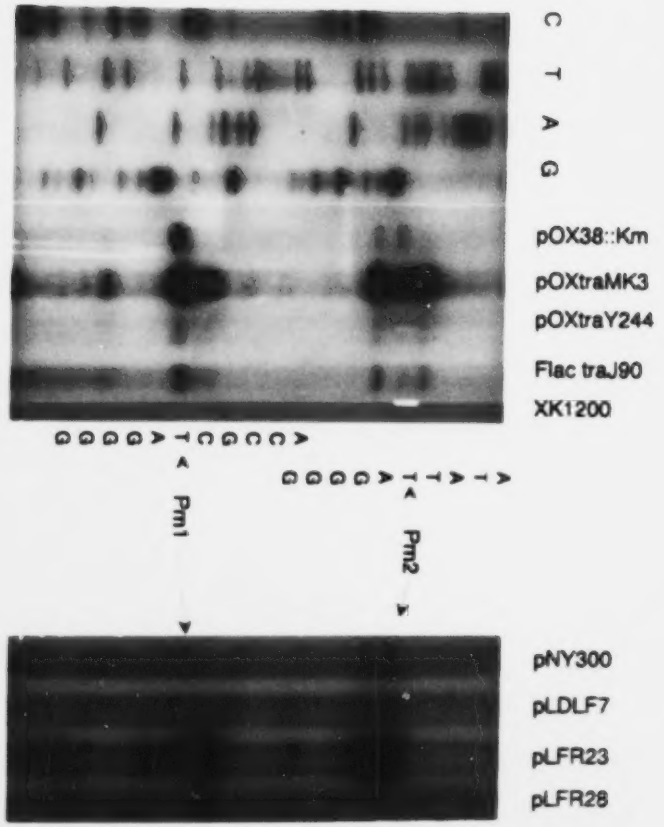

C
T
A
G
pOX38::Km
pOXtraMK3
pOXtraY244
Flac traJ90
XK1200
G G G G A T C G C C A G G G G A T T A T A
Pm1
Pm2
pNY300
pLDLF7
pLFR23
pLFR28

Figure 3.4

A schematic diagram showing the structure of the multicopy plasmids containing cloned fragments of the *oriT* region and *traM*. Protein binding sites within *oriT* are indicated, as well as the two *traM* promoters and the *oriT* site. The *traM* translation termination codon is indicated by an asterisk. The first line indicates the size in kilobases, while the third line represents the position of restriction enzyme sites within the *Bgl*II fragment. The vector used for the construction of plasmids shown here was pUC18.

1 KB
*
Pm2Pm1
sbmA
sbmB
IHFB
sbmC
sbyA
IHFA
oriT
BglII
DraI
SalI
DraI
BglII
pNY300
pLDLF7
pLFR23
pLFR28

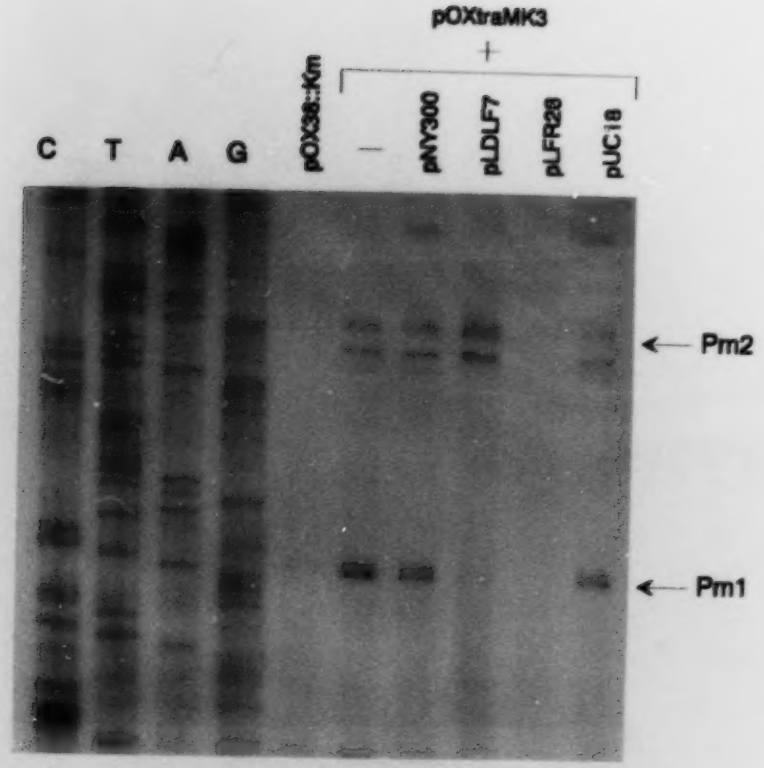
that were identified for the pOX38 plasmids were detected. Similarly, plasmid pLDLF7 (Di Laurenzio *et al.*, 1992), which lacks the region from the *Bgl*II site to *sbyA*, showed *traM* transcripts arising from both Pm1 and Pm2. Plasmid pLFR23 appears to express traM in an unregulated manner as transcripts arising from Pm1 and Pm2 were at least 10-times more abundant in this construct than in either of the other two multicopy plasmids tested. These results confirmed that pNY300 and pLDLF7, containing cloned *oriT* regions, are suitable for studying *traM* transcription and that the expression of *traM* from these plasmids is under similar control to that in pOX38::Km. The lack of detection of initiation at Pm1 and Pm2 from pLFR28 confirms that *traM* expression in this construct is vector driven.

Primer extension analysis on RNA isolated from cells carrying the same pOX derivatives as described above was carried out using primer SPE11 (Figure 3.1). This resulted in the identification of transcripts initiating at the same sites as was detected with SPE8 (data not shown), although transcripts initiating at Pm1 were difficult to distinguish from labelled primer, due to close proximity of this primer to Pm1. Results using primer SPE12 failed to detect any pOX38::Km-specific transcription initiation upstream of this primer binding site.

The -10 region for Pm1 lies within *sbmA*, suggesting that this promoter may be regulated by TraM. This was investigated by performing primer extension analyses (Figure 3.5) on RNA isolated from cells containing pOX*traM*K3 and one of pNY300, pLDLF7 or pLFR28, which all produce enough TraM to support transfer by pOX*traM*K3. TraM expressed from pLFR28, which lacks any of the *traM* promoters, completely repressed transcription from Pm1 and Pm2 of pOX*traM*K3. In the presence of of plasmids pNY300 and pLDLF7, some expression was detected from Pm1 and Pm2. One can conclude then that *traM* expression detected in these cells originates from *traM* promoters on these multicopy plasmids. A control lane showing the results of primer extension analysis on cells containing pOX*traM*K3 and pUC18, the vector used in the construction of the multicopy constructs, showed that this plasmid has no effect on *traM* expression. These results confirm that *traM* expression from Pm1 and Pm2 is negatively controlled by TraM. The relative intensities of bands at Pm1 and Pm2 in cells containing pOX*traM*K3 and pLDLF7 suggests that TraM binds Pm1 preferentially in pLDLF7, such that expression from Pm1 is almost completely repressed, while that from Pm2 is slightly enhanced. This effect is not seen in cells containing pOX*traM*K3 and pNY300. In comparison to pNY300, plasmid pLDLF7 lacks the primary TraY binding site in *oriT*,

Figure 3.5

Autoregulatory activity of TraM. Primer extension analysis of RNA extracted from cells containing the plasmids indicated, using primer SPE8. The corresponding nucleotide sequence is shown on the left and the arrows represent the 5' ends of transcripts Pm1 and Pm2.


pOXtraMK3
+
C
T
A
G
pOX38::Km
—
pNY300
pLDLF7
pLFR28
pUC18
Pm2
Pm1

sbyA. It was reasoned that the absence of *sbyA* in pLDLF7 (and hence absence of TraY binding adjacent to *sbmC*) may be responsible for this effect. This was tested by investigating *traM* expression from pNY300 in the presence of pOX*traY*244 (Figure 3.6). A similar result was obtained in these cells, as expression from Pm1 was almost completely repressed, while expression from Pm2 was enhanced.

IV. The role of *finO* in *traM* expression

The suggestion that FinO, required for fertility inhibition, represses *traM* expression (Dempsey, 1993) prompted me to investigate this. Initially, *finO* was expressed from plasmid pSNO104 (Lee *et al.*, 1992), which contains a 4.0 kb *Pst*I fragment that includes genes *traX*, *orf286*, *finO*, and *orfB*, which expresses a 12kDa protein of unknown function. Cells containing pACYC184 (Chang and Cohen, 1978), the vector used to express FinO, together with pOX38::Km, showed the presence of a 510 nucleotide transcript (Figure 3.7). This corresponds to a transcript which initiates at Pm1 and terminates within the region predicted to be a rho-independent transcription termination site (Thompson and Taylor, 1982). An identical transcript was present in pOX*traY*244, but at lower levels. In cells containing pOX38::Km and pSNO104, very little *traM* was detected by northern blot analysis. Similar results were obtained for primer extension analysis - the presence of pSNO104 in cells expressing *traM* from pOX38::Km resulted in a decrease of *traM* expression from Pm1 and Pm2. To determine whether this effect was specific to pOX38::Km, or whether FinO also affected *traM* expression on the F plasmid itself, immunoblot analysis was done and the results are shown in Figure 3.8. The presence of FinO in cells expressing *traM* from F or pOX38::Km had the same effect - TraM protein production was severely diminished. Since pSNO104 expressed a number of genes aside from *finO*, it could not be determined with certainty that this effect on *traM* transcription was mediated by FinO. Plasmid pTVB6.11 produces a truncated ORF286 protein, and increased amounts of FinO protein when compared with a similar plasmid containing a wild type *orf286* gene (van Biesen and Frost, 1992). The presence of this plasmid in cells also containing pOX38::Km resulted in barely detectable levels of TraM (Figure 3.9). In plasmids which produced moderate amounts of FinO due to a complete absence of *orf286* transcripts (pCB010 and pTVB2.1), a small reduction in TraM production was observed when compared to a control containing the vector alone (pUC118). Since *orf286* transcription has been demonstrated to be required for maximal FinO expression, these results suggest that FinO does decrease *traM* expression. The possibility of the 12 kDa protein encoded downstream of FinO playing a role in *traM* expression cannot, however, be ruled out at this stage.

Figure 3.6

The role of TraY in *traM* expression. Primer extension analysis on RNA extracted from cells containing the plasmids indicated, using primer SPE8. The corresponding nucleotide sequence indicated on the left was determined using pNY300 DNA as a template and the arrows represent the 5' ends of transcripts Pm1 and Pm2.

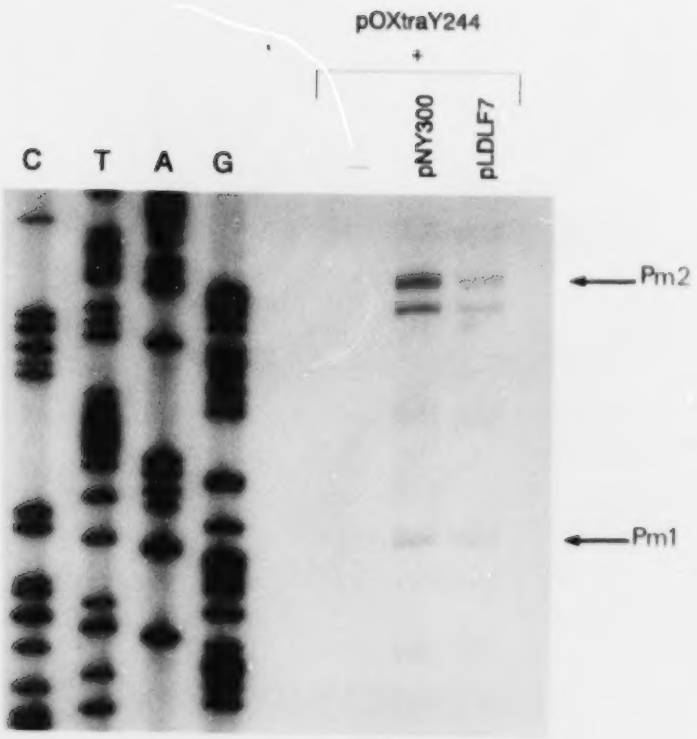

pOXtraY244
+
C
T
A
G
pNY300
pLDLF7
Pm2
Pm1

Figure 3.7

FinO decreases *traM* transcription from pOX38::Km. Northern blot analysis was performed on *traM* transcripts expressed from pOX38::Km in the absence (+pACYC184) or presence (+pSNO104) of FinO, and from the *traY* mutant, pOX*traY*244. The size of RNA molecular weight markers is given in kilobases.

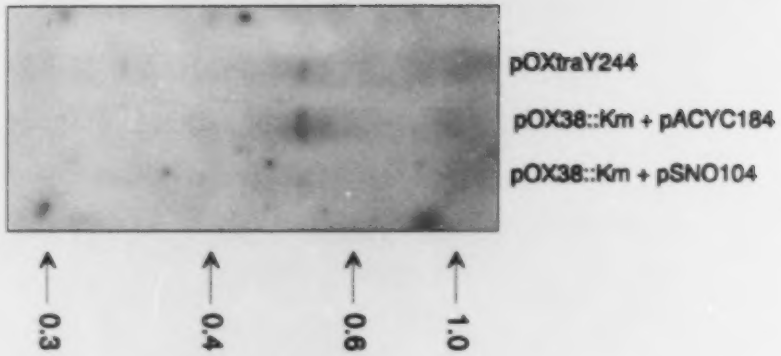

pOXtraY244
pOX38::Km + pACYC184
pOX38::Km + pSNO104
0.3
0.4
0.6
1.0

Figure 3.8

FinO decreases TraM production from F and pOX38::Km. Crude cell lysates were separated by SDS-PAGE and subjected to western blot anlaysis using polyconal TraM antisera. TraM production from F (M176) and pOX38::Km was monitored in the presence of vector alone (+pACYC184), or vector expressing functional FinO (+pSNO104).

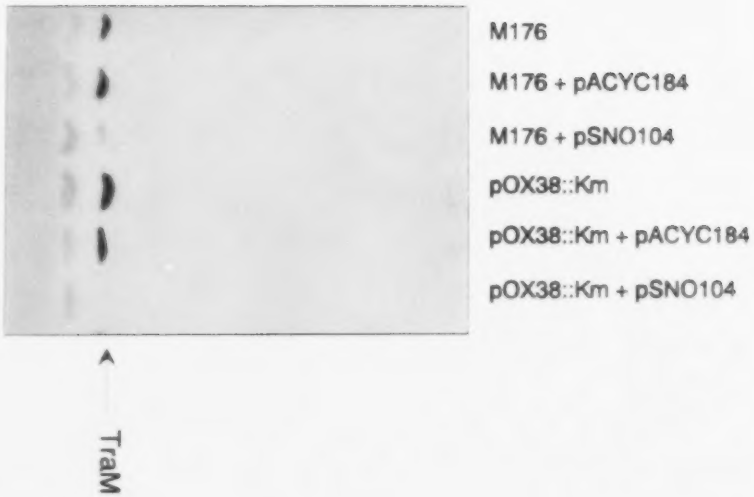
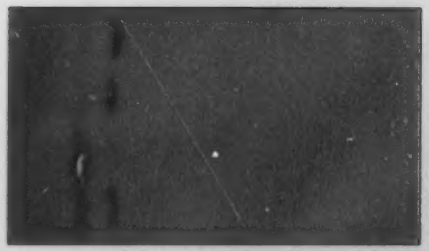

M176
M176 + pACYC184
M176 + pSNO104
pOX38::Km
pOX38::Km + pACYC184
pOX38::Km + pSNO104
TraM

Figure 3.9

FinO concentration affects TraM expression. Crude cell lysates were separated by SDS-PAGE and subjected to western blot analysis using polyconal TraM antisera. TraM production from pOX38::Km was monitored in the presence of vector alone (pUC118), vector expressing moderate amounts of FinO (pCB010 and pTVB2.1) and vector expressing large amounts of FinO (pTVB6.11).


pOX38::Km + pUC118
pOX38::Km + pCP010
pOX38::Km + pTVB6.11
pOX38::Km + pTVB2.1
TraM

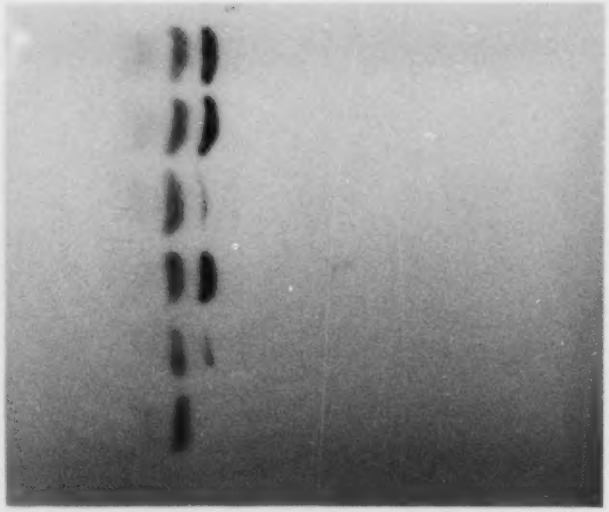
V. Analysis of TraM protein production

The number of TraM molecules in a cell containing the F plasmid was calculated by comparison of the intensity of a band generated by immunoblot analysis from a known number of cells (determined from the wet weight of the cells), with the intensity of a band obtained on immunoblot analysis of pure protein of known concentration. M176 cells were calculated to contain 27 molecules of TraM per cell, confirming the tight regulation of the strong *traM* promoters. Plasmids used in primer extension analyses were investigated for the amount of TraM that was produced to determine whether there was a correlation between the amount of protein produced and levels of *traM* transcription. Figure 3.10 shows the results of an immunoblot using crude lysates of cells carrying single copy plasmids expressing *traM*. Densitometric analysis of the intensities of the bands in each lane showed that *traJ* had no effect on TraM production, as an F *traJ* mutant expressed the same amount of TraM as wild type F*lac*. In contrast, there was a 10-fold reduction in the amount of TraM produced from a *traY* mutant, pOX*traY*244. Surprisingly, there also appeared to be a 10-fold reduction of TraM in the IHF$^-$ host, MC253. Since little effect on *traM* transcription in this host was detected, this implies a role for IHF at the translational level. As expected, no TraM was detected in cells carrying pOX*traM*K3.

Expression of TraM from the multicopy constructs used in this study was also investigated (data not shown). Plasmids pNY300 and pLDLF7 produced approximately the same amount of TraM, suggesting that the removal of the nick site does not affect TraM expression. However, pLFR23 produces at least 20-fold more TraM than either of the other two plasmids. This has recently been demonstrated to be the result of an mutation which causes a change from alanine to valine at amino acid 37, which lies within a motif that is remarkably non-conserved amongst F-like plasmids. Plasmid pLFR28, which lacks any of the *traM* promoters detected by primer extension analysis, produced approximately the same amount of TraM as pOX38::Km, which explains why it is able to complement pOX*traM*K3.

Figure 3.10

A comparison of TraM production from a variety of *tra*⁻ mutants. Crude cell lysates from an equal number of cells in each case were separated by SDS-PAGE and subjected to western blot analysis using polyclonal TraM antisera. Host strains for F *lac* and F *lac traJ* were JC3272. All other host strains were XK1200, except for pOX38::Km in MC253 which is indicated.


Flac
FlactraJ
pOX(MC253)
pOX38::Km
pOXtraY244
pOXtraMK3
6,2
TraM
12,3
18,4

C. Discussion

In 1972, a *tra*⁻ mutant of the F plasmid, M102, was described and the mutation was predicted to be a frameshift mutation within the *traM* gene (Achtman, *et al.*, 1972). The mutation has now been identified as an insertion of a G residue at nt 541 (Figure 3.1) and results in the production of a truncated protein of 44 amino acids (Frost, 1994). This mutant however, is not stable, and reverts to a wild type phenotype with high frequency. In order to study the effect of the *traM* gene product on *traM* transcription, an insertion mutant was constructed by interrupting the gene at the *Sal*1 site with a kanamycin resistance cassette. The inability of this *traM* insertion mutant, pOX*traM*K3, to transfer DNA despite the ability to synthesize and express F pili, confirms the requirement for a functional *traM* gene in conjugative DNA processing events (Kingsman and Willettts, 1972; Achtman *et al.*,1972). This requirement was further confirmed by the restoration of transfer to cells containing pOX*traM*K3 which were expressing TraM in *trans* from a second, compatible plasmid.

Previously, using *in vitro* run-off transcription techniques with linear templates (Thompson and Taylor, 1982), a promoter for the F *traM* gene was identified 165 bp upstream of the *Sal*1 site within the *traM* gene. Di Laurenzio *et al.* (1992) have suggested that the binding characteristics of TraM to the *oriT* region, where the binding sites are in phase at greater than 10.5 residues per turn, correspond to those of proteins which could alter the supercoiling within a constrained domain of DNA. The presence of negatively supercoiled templates has also been shown to be important in the efficiency of nicking by TraI (Reygers *et al.*, 1991; Matson and Morton 1991) and for transcription from the *traY* promoter (Gaudin and Silverman, 1993), indicating that the conformation of DNA is an additional regulatory factor involved in DNA transfer. This work describes the results of a detailed transcriptional analysis of F *traM* from native and recombinant supercoiled plasmids, in an effort to identify which factors were important in regulating the transcription of this gene.

Primer extension analysis was done on RNA from cells containing either multicopy plasmids containing cloned copies of the *oriT* region, or pOX38::Km, a low copy number transfer-proficient derivative of F, or the *traY* and *traM* pOX38 mutants, pOX*traY*244 and pOX*traM*K3, respectively. Except for pLFR23, the number and size of transcripts identified from the cloned *traM* genes containing the *traM* promoters corresponded to those found for the pOX38::Km plasmid, indicating that expression of *traM* from the

recombinant plasmids was regulated in the same way as on the native plasmid. However, due to the increased copy number of the cloned *traM* genes, these transcripts were more easily detected.

A major transcript corresponding to that identified previously (Thompson and Taylor, 1982) was expressed by all *traM*-containing plasmids. A second transcript initiating within the -35 region of Pm1 was also identified. This region contains a 22bp sequence that is an imperfect direct repeat of a sequence at Pm1 (Thompson and Taylor, 1982) and this second promoter, Pm2, initiated at the start of this repeat.

Expression from Pm1 and Pm2 of pOX38::Km was highly enhanced in the *traM* insertion mutant. The binding of F TraM within the *traM* promoter region has led to the suggestion that TraM regulates its own expression. Indeed, for the resistance plasmid R1, using *traM'lacZ* fusions present in the chromosome, it was demonstrated that this is the case (Schwab *et al.*, 1993), although it is not clear whether this effect applies to only one or both *traM* promoters. More recently, Abo and Ohtsubo (1993) have demonstrated by primer extension analysis and by measuring β-galactosidase activity from an R100 *traM-lacZ* fusion on a plasmid, that one of two promoters identified by them is negatively regulated by TraM expressed both in *cis* and in *trans*. A third *traM* promoter located within the *traM* ORF has been proposed for R100, but does not appear to be regulated by TraM (Dempsey, 1989). Our results confirm the autoregulatory nature of *traM* expression at both F *traM* promoters, but contrasts to the findings for R100, where the autoregulatory effect was seen only on Pm2. Since Pm1 and Pm2 both coincide with regions shown to be bound by TraM (Di Laurenzio *et al.*, 1992), one would expect that transcription from these promoters could be affected by this protein. In R100, the promoter that appeared to be constitutively expressed lies downstream of the four TraM binding sites (Abo and Ohtsubo, 1993), so binding of TraM to these sites would not be expected to preclude the binding of RNA polymerase to this promoter.

The expression of F *traM* in *trans* from a compatible multicopy plasmid repressed initiation from Pm1 and Pm2 of pOX*traM*K3 to that of wild type levels. The selective repression of Pm1 over Pm2 observed when pOX*traM*K3 was complemented with pLDLF7 rather than pNY300, suggested that additional factors within *oriT* determine which *traM* binding sites will be occupied. A similar result was obtained in cells containing the *traY* mutant, pOX*traY*244 and pNY300, suggesting that the binding of TraY to *sbyA* is responsible for derepression of expression at Pm1.

IHF has been shown to be required for conjugal transfer in both F and R100. However, the information available on the effect of IHF on *traM* transcription is contradictory. In one study, expression from the R100 *traM* promoter was decreased in an IHF$^-$ mutant (Dempsey and Fee, 1990), while in another (Abo and Ohtsubo, 1993), the expression from the same promoter was repressed to 60% of the fully expressed level by IHF. In the current study, while no decrease in *traM* transcription was detected in the absence of IHF, but a large decrease in the amount of TraM protein produced was observed. The nature of regulation of *traM* by IHF is clearly not fully understood and requires further investigation.

The intensity of *traM* transcripts detected by northern blot analysis was low, suggesting that very little *traM* mRNA is present. This could be the result of stringent regulation of the *traM* promoters, or instability of the mRNA. The lack of detection of transcripts corresponding to intiation at Pm2 could be due to the low level of expression from this promoter and poor resolution of transcripts differing by only 46 nucleotides in size. We investigated the effect of the FinO protein on *traM* expression. It has been suggested (Dempsey, 1993) that in R100, FinO inhibits *traM* expression by binding to the *traM* promoter. In work done in this laboratory, no binding of FinO to *oriT* DNA could be demonstrated (T. van Biesen, unpublished observations). In fact, recent work has shown that FinO is an RNA binding protein, which binds to an RNA stem of 14 base pairs or more (Van Biesen and Frost, 1994). We have shown here that the number of *traM* transcripts is markedly decreased in cells which also express FinO. It is currently being investigated whether this effect is due to a direct interaction between FinO and *traM*, or whether the effect is mediated through other plasmid- or host proteins. The presence of a 17 bp stem in the *oriT* region, located at *sbmC*, suggested that this may be the site of action of FinO in *oriT*, but no evidence has been found that this region of *oriT* is transcribed. Also, it is not clear at this stage how the binding of FinO to *sbmC* would affect *traM* transcription, or if the effect is mediated through RNA stability.

The picture that has emerged on the regulation of *traM* is complex, with a number of host- and plasmid-encoded factors controlling the expression of this protein, both at the transcriptional and translational level. Clearly, the understanding of *traM* expression requires an understanding of how the various proteins interact with *oriT* DNA and with each other.

D. Materials and methods

I. Bacterial strains, plasmids and media

Bacterial strains used are described in Table 3.2. The recipient used in triparental mating experiments, *E. coli* XK1200, was a gift from K. Ippen-Ihler, as were plasmids pOX38, pOX38::Km and pOX*traY*244. pOX38 (Guyer *et al.*, 1980) is a transfer proficient derivative of the F plasmid, containing the entire transfer region within a 55kb *Hin*dIII fragment, while pOX38::Km (Chandler and Galas, 1983) also contains a kanamycin resistance cassette inserted at the *Hin*dIII site. pOX*traY*244 was constructed by inserting a kanamycin resistance gene between the *BclI* sites in the second half of *traY* in pOX38 (K. Ippen-Ihler, personal communication) and the construction of a plasmid containing a *traM* insertion mutant is described below. Plasmids pLFR23 and pLFR28 were constructed using pNY300 as a template for the PCR with primers 5'-CACCGTTTTGTAGGGGTGGTAC-3' and 5'-CGTCCCTGTTTGCATTATGA-3' respectively, and the Universal primer which binds within the vector sequence of pNY300. All other plasmids have been described previously and are shown in Table 3.3. *E.coli* cells were grown in LB (Luria-Bertani) medium (Maniatis *et al.*, 1982) containing the appropriate antibiotics at the following concentrations: ampicillin ($50\mu g.ml^{-1}$), kanamycin ($25\mu g.ml^{-1}$), chloramphenicol ($50\mu g.ml^{-1}$), and naladixic acid ($40\mu g.ml^{-1}$).

II. Recombinant DNA techniques and reagents

Restriction and modification enzymes and dNTPs were purchased from Boehringer-Mannheim and, unless otherwise stated, used according to the manufacturer's instructions. The kanamycin resistance cassette (Kan^r GenBlock) used to create the *traM* insertion mutant was purchased from Pharmacia, as was RNAguard. Sequencing was done using the dideoxy chain termination method (Sanger *et al.*, 1977) and the reagents were purchased from New England Biolabs or United States Biochemicals. DNA was isolated from agarose gels by centrifuging the excised fragment through silanized glass wool at 12000g for 20 minutes and recovering the eluate containing the DNA. This was extracted with phenol and the DNA was precipitated with ethanol.

III. Creation of a *traM* insertion mutant

Plasmid pLDLF7 (Di Laurenzio *et al.*, 1992) was constructed by cloning the *DraI* fragment encoding the entire *traM* gene (Figure 3.4) into pUC18 (Vieira and Messing,

Table 3.2

Description of bacterial strains used in the this study

Bacterial strain	Genotype	Source/Reference
JC3272	F⁻, *lacX174, gal, his, trp, lys, rpsL, tsx*	Achtman, *et al.*, (1971)
M176	F *lac*	Achtman, *et al.*, (1971)
XK1200	F-, *lacU124,* (*nadA, gal, att, bio*) *gyrA*	Moore, *et al.*, (1987)
RD17	(*pro-lac*)XIII, ⁻, *recA56, rel1, supE44, thi-1*	Tsai , *et al.*, (1987)
MC253	*ara,* Δ(*lac pro*) *gyr,A metB, argE, rif, thi, supF* Δ 82(*himA*) Δ3 (*hip*)*:cam*	Gamas, *et al.*, (1986)

Table 3.3

Description of plasmids used in this study

Plasmid	Relevant Genotype	Source/Reference
pLDLF7	*traM*[a]	Di Laurenzio *et al.*, (1992)
pNY300	*oriT, traM, finP*[b]	Frost *et al.*, (1989)
pMF7-Kan	pLDLF7 with *traM* interrupted by KanR	This work
pSNO104	FinO+[c]	Lee *et al.*, (1992)
pRS27	*oriT, traM, finP, traJ, A, L, E, K, B*[d]	Skurray *et al.*, (1978)
pOX38	*tra*$^{+}$[e]	Guyer *et al.*, (1980)
pOX38::Km	*tra*$^{+}$, KanR[f]	Chandler and Galas (1983)
pOX*traY*244	*tra*$^{-}$, KanR, TraY-	K. Ippen-Ihler
pOX*traM*K3	*tra*$^{-}$, KanR, TraM-	This work
F*lac*	*tra*$^{+}$, *lac*$^{+}$	Achtman *et al.*, (1971)
F*lac traJ*90	*tra*$^{-}$, *lac*$^{+}$	Achtman *et al.*, (1971)

a 680bp *Dra*I fragment from F inserted into pUC18 such that *traM* is expressed from its own promoter.
b 1.1kb *Bgl*II fragment shown in figure1 inserted into the *Bam*H1 site of pUC18 such that *traM* is expressed from its own promoter.
c 4kb *Pst*I fragment containing *orf286, finO* and a 12kDa protein in pACYC184 (Chang and Cohen, 1978).
d Contains the 9kb *f6* fragment inserted in pSC101.
e 55.4kb *Hin*dIII fragment of F, containing the whole *tra* region, recircularized.
f As above, but also contains a kanamycin resistance gene inserted at the *Hin*dIII site.

1982). The insert was excised by digestion with *Hin*dIII and *Eco*RI and gel purified. This fragment was subsequently digested with *Sal*I and ligated to the *Sal*I digested Kan^r cassette in a reaction mixture which also contained *Hin*dIII and *Eco*RI digested pUC18. This resulted in the generation of a plasmid identical to pLDLF7, except that the *traM* gene was interrupted with the Kan^r cassette. Plasmid pMF7-Kan contained a 2020 bp insert, comprising the 680 bp *Dra*I fragment and the 1.34 kb kanamycin resistance cassette, inserted at the *Sal*I site within the *traM* gene.

IV.Triparental mating to create pOX*traM*K3

Donor (*E. coli* RD17/pOX38) and recipient (*E. coli* JC3272/pMF7-Kan) cells were grown to mid-log phase and an equal volume (0.1ml) of each was mixed in 1ml of LB for 1 hour at 37^{o}C. Subsequently, 0.4ml of a mid-log phase culture of XK1200 was added and incubated for a further 6 hours. Transconjugants were selected on plates containing kanamycin and naladixic acid. All transconjugants were tested for their ability to express F pili using a f1 phage sensitivity spot test. In addition, plasmid DNA from f1 sensitive transconjugants was extracted by the alkaline lysis method (Birnboim and Doly, 1979) and examined by gel electrophoresis and unblot hybridization analysis (Sambrook *et al.*, 1989) to confirm the presence of the 1.3 kb Kan^r cassette within the 1.1 kb *Bgl*II fragment of pOX38.

V. Mating efficiency of pOX*traM*K3

Mating assays were performed as described previously (Frost *et al.*, 1989) and the recipient used was JC3272. Briefly, equal volumes (0.1ml) of mid-log phase donor and recipient cells were mixed in 1ml of LB and incubated for 30 minutes at 37^{o}C. The mating mixture was vortexed to disrupt pili and prevent further transfer, and 10µl volumes were spotted onto plates containing appropriate antibiotics which selected for transconjugants. The number of donors in the culture was estimated by selecting for donor cells from the mating mixture with appropriate anitibiotics. Plasmid pOX 38::Km, containing the kanamycin resistance cassette outside of the transfer region, was used as a positive control and mating efficiencies of all other plasmids were compared to this. Plasmids pNY300, pLDLF7 and pLFR28 were used as a source of *traM* in assays to assess complementation of the insertion mutant. The ability of the *traY* mutation to be complemented in *trans* was assayed using pRS27 as a source of *traY*.

VI. Primer extension analysis

Three synthetic oligonucleotide primers were generated for primer extensions: SPE8 (5'-CATAGGCATCATTGCTGATATACAG-3') bound 118 nt downstream from the previously reported *traM* promoter; SPE11 (5'-CATAATGCAAACAGGGACGCACCG-3'), bound 16 nt downstream of the previously reported *traM* promoter; and SPE12 (5'-CACACCGCGCCGCTAGCAGC-3'), bound 21 nt upstream of the reported *traM* promoter (Figure 3.1). When analysing RNA from cells containing *traM* on a single or low copy number plasmid, 30µg of RNA was allowed to anneal to approximately 0.5 pmol of primer, labelled at the 5'end with [γ^{32}P]-ATP using T4 polynucleotide kinase (Sambrook *et al.*, 1989). The primer and RNA were mixed in a 30µl volume of buffer containing 3M NaCl, 0.5M Tris-HCl, pH7.5 and 1mM EDTA, pH8.0, denatured at 85°C for 5 minutes, then allowed to anneal at 37°C for a minimum of 1 hour. Following annealing, the nucleic acids were precipitated with ethanol and the pellets allowed to air dry. They were resuspended in a 25µl volume of AMV reverse transcriptase buffer containing 0.5mM dNTP's and 15 units RNAguard. AMV reverse transcriptase (20units) was added and the reaction was incubated at 42°C for 1 hour. RNA was removed by treatment with RNaseA for 15 minutes at 37°C and the DNA was precipitated with ethanol in the presence of 0.3M NaOAc and 10µg glycogen. The products were separated on a 6% denaturing polyacrylamide gel alongside dideoxy sequencing reactions using the same primer as was used for the extension. When analysing RNA from cells containing multicopy plasmids expressing *traM*, the procedure was identical except that only 10µg of RNA was used.

VII. RNA preparation

RNA was isolated using the modified hot phenol extraction method described previously (Frost *et al.*, 1989). When necessary, RNA was treated with RNase-free DNaseI at room temperature for 1 hour, then phenol extracted, precipitated with ethanol and the pellet was resuspended in water.

VIII. Northern blot analysis

Oligonucloetide SPE8 (50pmol) was end-labelled with [γ^{32}P]-ATP using T4 polynucleotide kinase (Sambrook *et al.*, 1989) and used as a probe for *traM* transcripts. RNA (20µg) was boiled prior to separation by electrophoresis on an 8% polyacrylamide gel containing 8M urea, and transferred to ZetaProbe nylon membrane (Biorad) using a Biorad Trans-Blot cell. The membrane was prehybridized for a minimum of 2 hours at 37°C in 2.5 x SSC, 5 x Denhardt's solution (Sambrook *et al.*, 1989), 1.5% SDS, 100 µg.ml^{-1} of *E. coli* strain W tRNA type XX (Sigma) and 100 µg.ml^{-1} of denatured calf

thymus DNA. Hybridization in the presence of 10^6 cpm/ml hybridization solution was at 37°C overnight and the membrane was washed twice for 15 minutes each time with 6 x SSC, 0.1% SDS at room temperature. Autoradiography was performed at -70°C with an intensifying screen using Kodak X-AR5 film.

IX. Analysis of TraM protein production

For analysis of crude protein lysates, cells were grown to an optical density of approximately 0.8. Equal numbers of cells (as determined by optical density values) were pelleted and boiled in cracking buffer containing 60mM Tris-HCl, pH6.8, 1% SDS, 10% glycerol, 1% β-mercaptoethanol and 0.02% bromophenol blue prior to separation on a 15% SDS-PAGE gel. Proteins were transferred to Immobillon nylon membranes (Millipore) using a Trans-blot Semi-Dry Transfer Cell (Biorad). Conditions for transfer were those recommended by the supplier (3.5 mA/cm^2) and transfer was allowed to proceed for 30 minutes (mini-gels) and 45 minutes (large gels). Blocking reagent used to mimimize non-specific cross-reactivity was 10% skim milk powder (Difco) and the blocking reaction was carried out for 1 hour at room temperature. Immunological detection was done using polyclonal antisera raised against purified TraM and an ECL Detection Kit (Amersham), used according to the manufacturer's instructions.

E. References

Abo, T., and Ohtsubo, E. (1993) Repression of the *traM* gene of plasmid R100 by its own product and integration host factor at one of the two promoters. *J. Bacteriol.* **175:** 4466-4474.

Achtman, M., Willetts, N.S. and Clark, A.J. (1971) Beginning a genetic analysis of conjugational transfer determined by the F factor in *Escherichia coli* by isolation and characterization of transfer-deficient mutants. *J. Bacteriol.* **106:** 529-538.

Achtman, M., Willetts, N., and Clark, A.J. (1972) Conjugational complementational analysis of transfer-deficient mutants of F*lac* in *Escherichia coli*. *J. Bacteriol.* **110:** 831-842.

Birnboim H. C., and J. Doly. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7:** 1513-1523.

Bramhill, D., and Kornberg, A. (1988) A model for initiation at origins of DNA replication. *Cell* **54:** 915-918.

Chandler, M., and Galas, D. (1983) Cointegrate formation mediated by Tn9 II. Activity of *IS*1 is modulated by external DNA sequences. *J. Mol. Biol.* **170:** 61-91.

Chang, A. C. Y., and Cohen, S.N. (1978) Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J. Bacteriol.* **134:** 1141-1156.

Dempsey, W.B. (1989) Sense and antisense transcripts of *traM*, a conjugal transfer gene of the antibiotic resistance plasmid R100. *Mol. Microbiol.* **3:** 561-570.

Dempsey, W.B. (1993) Key regulatory aspects of transfer of F-related plasmids. *In* Bacterial Conjugation, p. 53-73. D. B. Clewell, (ed). New York, New York: Plenum Press.

Dempsey, W.B., and Fee, B.E. (1990) Integration host factor affects the expression of two genes at the conjugal transfer origin of plasmid R100. *Mol. Microbiol.* **4:** 1019-1028.

Di Laurenzio, L., Frost, L.S., Finlay, B.B., and Paranchych, W. (1991) Characterization of the *oriT* region of the IncFV plasmid pED208. *Mol. Microbiol.* **5:** 1779-1790.

Di Laurenzio, L., Frost, L.S., and Paranchych, W. (1992) The TraM protein of the conjugative plasmid F binds to the origin of transfer of the F and ColE1 plasmids. *Mol. Microbiol.* **6:** 2951-2959.

Friedman, D. I. (1988) Integration host factor: a protein for all reasons. *Cell* **55:** 545-554.

Frost, L. S. (1994). Unpublished observations.

Frost, L. S., K. Ippen-Ihler, and R. A. Skurray. (1994) Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol. Rev.* **58:** 162-210.

Frost, L.S., Lee, S., Yanchar, N., and Paranchych, W. (1989) *finP* and *fisO* mutations in FinP anti-sense RNA suggest a model for FinOP action in the repression of bacterial conjugation by the F plasmid. *Mol. Gen. Genet.* **218:** 152-160.

Gaffney, D., Skurray, R., and Willetts, N. (1983) Regulation of the F conjugation genes studied by hybridization and *tra-lacZ* fusion. *J. Mol. Biol.* **168:** 103-122.

Gamas, P., A. -C. Burger, G. Churchward, L. Caro, D. Galas, and M. Chandler. (1986) Replication of pSC101: effects of mutations in the *E. coli* DNA binding protein IHF. *Mol. Gen. Genet.* **204:** 85-89.

Gaudin, H.M., and Silverman, P.M. (1993) Contributions of promoter context and structure to regulated expresssion of the F plasmid *traY* promoter in *Escherichia coli* K-12. *Mol. Microbiol.* **8:** 335-342.

Guyer, M.S., Reed, R.R., Steitz, J.A., and Low, K.B. (1980) Identification of a sex-factor affinity site in *E. coli* as gamma-delta. *Cold Spring Harbor Symp. Quant. Biol.* **45:** 135-140.

Ippen-Ihler, K., and S. Maneewannakul. (1991) Conjugation among enteric bacteria: mating systems dependent on the expression of pili, p.35-39. *In* M. Dworkin (ed.), Microbial cell-cell interactions. American Society for Microbiology. Washington, D. C.

Kingsman, A., and Willetts, N. (1978) The requirements for conjugal DNA synthesis in the donor strain during the F*lac* transfer. *J. Mol. Biol.* **122:** 287-300.

Koronakis, V.E., Bauer, E., and Högenauer, G. (1985) The *traM* gene of the resistance plasmid R1: comparison with the corresponding sequence of the *Escherichia coli* F factor. *Gene* **36:** 79-86.

Lahue, E.E. and Matson, S. W. (1990) Purified *E. coli* F-factor TraY binds *oriT*. *J. Bacteriol.* **172:** 1385-1391.

Lee, S.J., Frost. L.S., and Paranchych W.P. (1992) FinOP repression of the F plasmid involves extension of the half-life of FinP antisense RNA by FinO. *Mol. Gen. Genet.* **235:** 131-139.

Luo, Y., Gao, Q., and Deonier, R.C. (1993) Mutational and physical analysis of F plasmid *traY* protein binding to *oriT*. *Mol. Microbiol.* **11:** 449-458

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual. (1st ed.) Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Matson, S.W., and Morton, B.S. (1991) *Escherichia coli* DNA helicase 1 catalyzes a site- and strand-specific nicking reaction at the F plasmid *oriT*. *J. Biol. Chem.* **266:** 16232-16237.

Moore, D., Wu, J.H., Kathir, P., Hamilton, C.M., and Ippen-Ihler, K. (1987) Analysis of transfer genes and gene products within the *traB-traC* region of the *E. coli* fertility factor, F. *J. Bacteriol.* **169:** 3994-4002.

Mullineaux, P., and N. Willetts. (1985) Promoters in the transfer region of plasmid F. *Basic Life Sci.* **30:** 605-614.

O'Connor, M.B., and Malamy, M.H. (1984) Role of the F-factor oriV1 region in *recA*-independent illegitimate recombination. Stable replicon fusions of the F derivative pOX38 and pBR322-related plasmids. *J. Mol. Biol.* **175:** 263.

Reygers, U., Wessel, R., Muller, H., and Hoffman-Berling, H. (1991) Endonuclease activity of *Escherichia coli* DNA helicase I directed against the transfer origin of the F factor. *EMBO J.* **10:** 2689-2694.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, (2nd ed.) Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74:** 5463-5467.

Schwab, M., Reisenzein, H., and Högenauer, G. (1993) TraM of plasmid R1 regulates its own expression. *Mol. Microbiol.* **7:** 795-803.

Skurray, R.A., Nagaishi, H., and Clark, A.J. (1978) Construction and *Bam*HI analysis of chimeric plasmids containing *Eco*RI DNA fragments of the F sex factor. *Plasmid* **1:** 174-186.

Thompson, R., Taylor, L., Kelly, K., Everett, R., and Willetts, N. (1984) The F plasmid origin of transfer: DNA sequence of wild-type and mutant origins and location of origin-specific nicks. *EMBO J.* **3:** 1175-1180.

Thompson, R., and Taylor, L. (1982) Promoter mapping and DNA sequencing of the F plasmid transfer genes *traM* and *traJ*. *Mol. Gen. Genet.* **188:** 513-518.

Tsai, M.-M., Fu, F.Y.-H., and Deonier, R.C. (1990) Intrinsic bends and integration host factor binding at F plasmid *oriT*. *J. Bacteriol.* **72:** 4603-4609.

Tsai, M.-M., Wong, R.Y.-P., Hoang, A., and Deonier, R. (1987) Transposition of Tn1000: in vivo properties. *J. Bacteriol.* **169:** 5556-5562.

van Biesen, T., and Frost, L.S. (1992) Differential levels of fertility inhibition among F-like plasmids are related to the cellular concentration of *finO* mRNA. *Mol. Microbiol.* **6:** 771-780.

van Biesen, T., and Frost, L.S. (1994) The FinO protein of IncF plasmids binds FinP antisense RNA and its target, *traJ* mRNA, and promotes duplex formation. *Mol. Microbiol.* In press

Vieira, J., and Messing, J. (1982) The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing using synthetic universal primers. *Gene* **19:** 259-268.

Willetts, N. S., and R. A. Skurray. (1987) Structure and function of the F factor and mechanism of conjugation, p. 1110-1133. *In* F. C. Niehardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umberger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D. C.

Willetts, N. S., and B. Wilkins. (1984) Processing of plasmid DNA during bacterial conjugation. *Microbiol. Rev.* **48:** 24-41.

Chapter 4

The role of the C-terminus of TraM in transfer: Characterization of a 3' deletion mutant of *traM*

A. Introduction

A comparison of the amino acid sequences of the 5 alleles of *traM* shows a high degree of sequence identity throughout the proteins, with most of the non-identical residues occuring in the amino terminal regions (Figure 4.1). Despite this sequence similarity between the F-like TraM proteins, TraM function appears to be allele specific, as only in a few instances can weak heterologous complementation be demonstrated (Willetts, 1981; Willetts and Maule, 1986; Goldner *et al.*, 1987; Schwab *et al.*, 1993). It has been suggested that the amino terminus of *traM* is the DNA-binding domain and that this region is responsible for the observed allele specificity of TraM binding (Koronakis *et al.*, 1985; Schwab *et al.*, 1993, Schwab *et al.*, 1994).

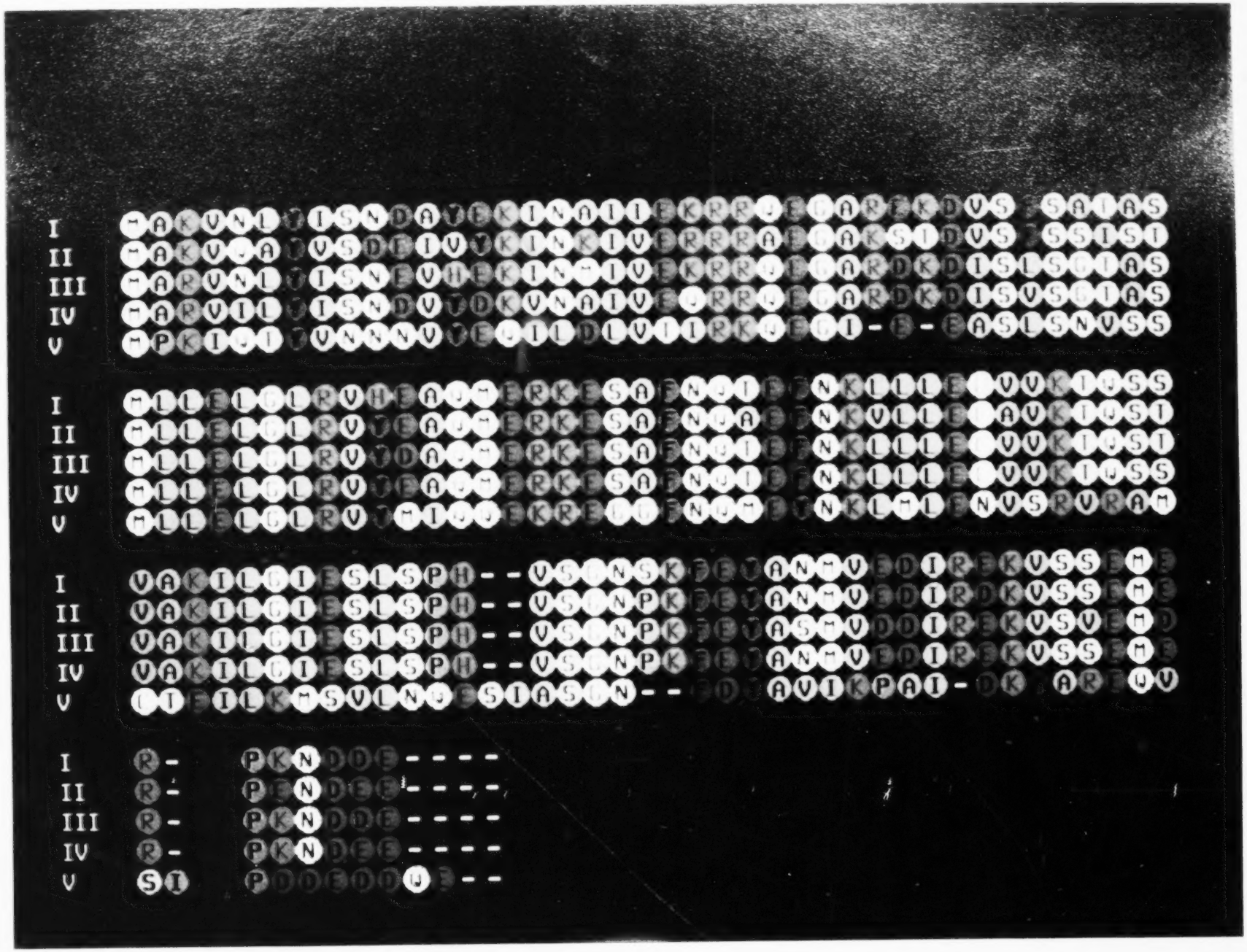
In the IncFII plasmid R1, the region encoding the first 22 amino acids can be folded into a near perfect amphiphilic alpha helix (Schwab *et al.*, 1993). Despite the amino acid differences between the various *traM* alleles in the N-terminus, all are capable of taking up this structure. Using site-specific mutagenesis of the N-terminus of the TraM protein of plasmid R1, Schwab *et al.* (1993) showed by measuring β-galactosidase activity that this region was involved in autorepression of a *traM-lacZ* fusion and in the binding to *oriT*, as judged by gel retardation analysis.

Inspection of the C-terminal regions of the various *traM* alleles reveals a highly conserved acidic tail featuring a preponderence of aspartic acid and glutamic acid residues, suggesting that this region may be critical for function. Surprisingly, the removal of the terminal third of the R1 TraM protein did not severely affect the ability of the truncated protein to repress a *traM-lacZ* fusion in *trans* (Schwab, 1993). It was therefore concluded that the N-terminal region of TraM was important for DNA-binding activity and autorepression. The function of the C-terminus was not investigated further. The domains of TraM involved in tetramer formation and the proposed signalling activity of the protein have not been identified.

This chapter describes the construction of a 3' terminal deletion mutant of F *traM* by the removal of 8 amino acids at the 3' terminus, in order to determine the function of the acidic tail of the protein. The C-terminal deleted protein produced by this construct was characterized according to its ability to support transfer, bind DNA, repress *traM* promoters Pm1 and Pm2, and form tetramers.

Figure 4.1

Amino acid sequence alignment of the 5 TraM alleles. Amino acids are reported using the one letter code and colour-coded according to the nature of the side chain, where blue=basic residues, red=acidic residues, green=hydrophobic residues, brown=aromatic residues, white=glycine, and yellow=polar residues. The sequences of the TraM protein of F (I), ColB4 (II), R1 (III), R100 (IV), and pED208 (V) plasmids are given. This figure was produced by Dr. Laura Frost and Dr. Brett Finlay, using a program designed by David Bacon, Department of Biochemistry, University of Alberta.


I
II
III
IV
V

B. Results

I. Construction of a C-terminal mutant, *traMΔ8*

In order to investigate the role of the C-terminus of TraM in transfer, a deletion mutant (pSPE2307) which lacked the terminal 8 amino acids of the protein (Figure 4.2) was constructed using PCR. Attempts to generate additional C-terminal mutants with fewer amino acids deleted have thus far been unsuccessful. This can partly be attributed to the high AT content of the 3' terminus of *traM*, which makes the design of specific PCR primers difficult. In addition, it appears that such products are lethal to cells when cloned, as all attempts at cloning resulted in the generation of vector deletions. Although a plasmid expressing wild type *traM* was already available (pLDLF7), a second plasmid was constructed (pSPE2309) which expressed wild type *traM* from a fragment with an identical 5' end to pSPE2307, in order that a true comparison of the expression of the wild type and mutant proteins could be made. A second series of plasmids expressing wild type *traM* and *traMΔ8* (pSPE2809 and pSPE2807 respectively) was constructed to test the ability of the mutant protein to repress *traM* expression *in trans* by primer extension analysis. Neither of these two plasmids contained any of the three TraM binding sites previously characterised (Di Laurenzio *et al.*, 1992) or the two *traM* promoters identified in Chapter 2. Fragments containing *traM* and *traMΔ8* genes were also cloned into the overexpression vector pT7.4 (Tabor and Richardson, 1985), to generate pSPE2309-7 and pSPE2307-7, respectively. *In vivo* labelling of overexpressed proteins demonstrated the production of a protein with an apparent MW on SDS-PAGE of 10,950 Da from pSPE2309-7, and a protein of apparent MW of 9,950 Da from pSPE2307-7 (data not shown).

II. Immunological detection of TraM

The expression of wild type and C-terminal deleted TraM was confirmed using polyclonal anti-TraM antisera and chemiluminescent detection. Figure 4.3 shows that pSPE2309 produced a protein which migrated with the same mobility on SDS-PAGE as wild type TraM produced by pLDLF7. A protein which migrated with slightly greater mobility than the wild type protein was detected in cells expressing TraMΔ8 from pSPE2307. Since equal numbers of cells were loaded for each sample and the bands are approximately of equal intensity, these results suggest that similar amounts of TraM are produced by these three constructs.

Figure 4.2

A schematic diagram of the plasmids used to express TraM and TraMΔ8. pSP2309 and pSP2307 contain both *traM* promoters and a wild type and 3' deleted (*traM*Δ8) *traM* gene respectively. pSP2809 and pSP2807 lack both *traM* promoters and contain a wild type and 3' deleted *traM* fragment respectively. The terminal 15 amino acids of TraM are shown, with the amino acids that were removed in the construction of the 3' deletions indicated in bold type. The asterisk denotes the translation termination codon of *traM*. Features of *oriT* are as described in Figure 1.6.

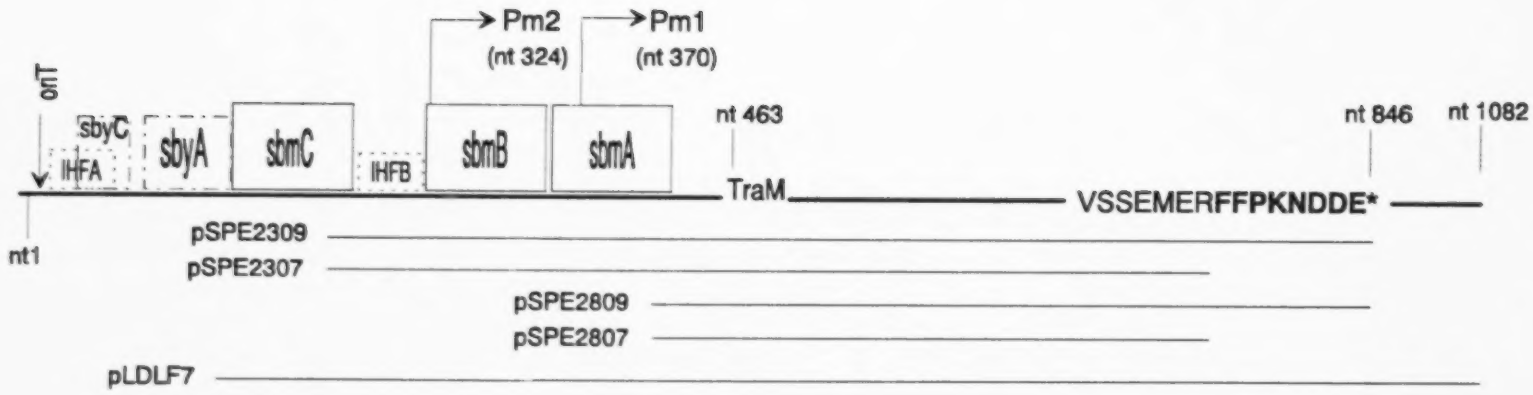
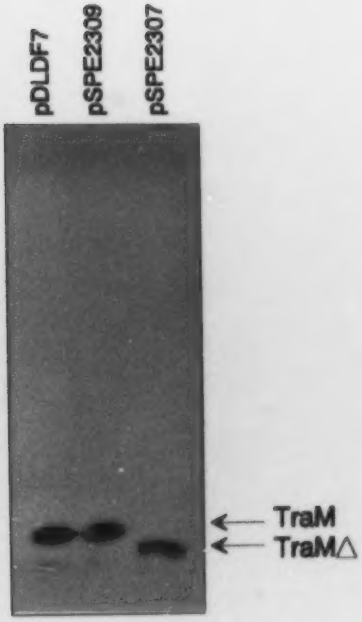

Pm2
(nt 324)
Pm1
(nt 370)
oriT
sbyC
IHFA
sbyA
sbmC
IHFB
sbmB
sbmA
nt 463
TraM
nt 846
nt 1082
VSSEMERFFPKNDDE*
nt1
pSPE2309
pSPE2307
pSPE2809
pSPE2807
pLDLF7

Figure 4.3

Western blot analysis of TraM and TraMΔ8 production. Plasmids containing wild type *traM* (pLDLF7 and pSPE2309) and 3' deleted (pSP2307) *traM* genes were used to express TraM and TraMΔ(8), respectively. Proteins from crude cell lysates were separated by SDS-PAGE and polyclonal TraM antisera was used to detect TraM.


pDLDF7
pSPE2309
pSPE2307
TraM
TraMΔ

III. Mating assays

The ability of TraMΔ8 expressed by pSPE2807 or pSPE2307 to complement pOX*traM*K3 was compared to that of similar plasmids expresssing wild type TraM (Table 4.1). Both plasmids pSPE2809 and pSP2309 were able to complement pOX*traM*K3 in mating assays. Plasmid pSPE2309 complemented transfer 5 times more efficiently than pSPE2809, probably as a result of the expression of more TraM from the former plasmid. TraMΔ8 expressed by pSPE2807 or pSPE2307 was unable to support transfer, indicating that the C-terminus of TraM is important for function.

IV. FPLC purification of crude extracts

Crude extracts from IPTG-induced cells bearing pSPE2309-7 and pSPE2307-7 were partially purified by gel exclusion chromatography. The propensity of TraM to aggregate upon purification (Schwab *et al.*, 1991; Di Laurenzio, 1992) made pure protein preparations unsuitable for gel filtration chromatography. Therefore, crude extracts were used to examine the quaternary structure of TraM and TraMΔ8. Analysis of every third fraction by immunoblot analysis showed that most of the wild type TraM was eluted from the column in fractions 35 - 66, with the bulk of the protein eluting in fractions 54 - 57 (Figure 4.4). The origin of the band in fractions 35 - 42 which migrates with slightly lower mobility than TraM is presently not known. The TraM-containing fractions (35-66) were pooled and used in gel retardation analyses. Similarly, TraMΔ8, expressed from pSPE2307-7 was eluted from the column in fractions 35 - 66, with the bulk of the protein eluting in fractions 54 - 57. TraMΔ8 in these fractions appeared to be slightly degraded as evidenced by smearing below the bands. Fractions 35 - 66 were pooled for use in gel retardation analysis.

A comparison of the elution profile of these two proteins with that of a series of standards allowed the calculation of the molecular weight of TraM and TraMΔ8. BSA (bovine serum albumin, MW 66 kDa) was eluted from the column in fractions 48 - 55 and peaked at fraction 51, while carbonic anhydrase (MW 29 kDa) was eluted in fractions 61 - 68, peaking at fraction 65. A plot of molecular weight vs Ve/Vo, where Ve is the elution volume of the protein and Vo is the void volume, which corresponds to the elution volume of a large molecule such as Blue Dextran (42 ml), was used to calculate the size of the native TraM protein (Figure 4.5). These results suggested that both TraM and TraMΔ8 occurred as a tetramer, with a calculated molecular weight of 50 kDa, which corresponds to the predicted size of 58 kDa for the tetramer. The appearance of both the wild type and mutant TraM proteins in the same fractions, suggested that the

Table 4.1

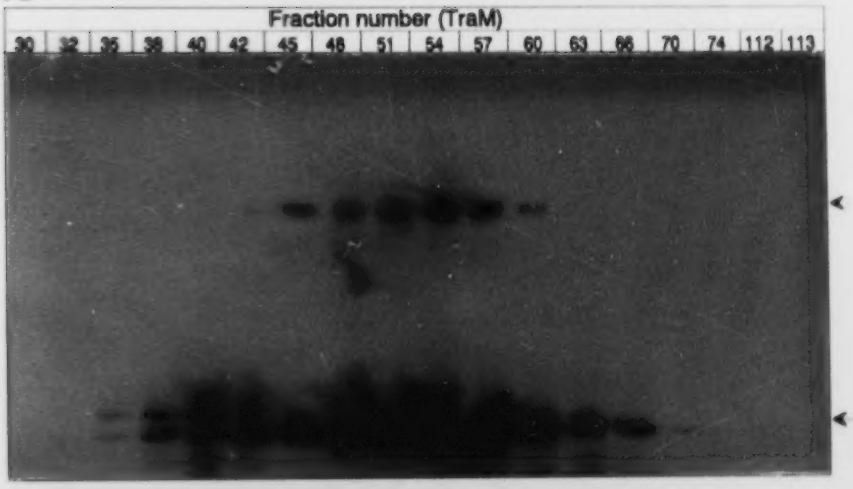
Mating efficiencies of cells expressing *traM*Δ8 compared with those expressing *traM*. Donor cells were allowed to mate with recipients and the number of transconjugants obtained per 100 donor cells is shown. Values shown are from a representative assay.

Donor	#Transconjugants/100Donors	Relevant Genotype
pOX*traM*K3	0	$TraM^-$
pOX*traM*K3 + pSPE2309	11	$TraM^+$ (Pm1 and Pm2)
pOX*traM*K3 + pSPE2307	0	TraMΔ8 (Pm1 and Pm2)
pOX*traM*K3 + pSPE2809	2	$TraM^+$ (ΔPm1 and Pm2)
pOX*traM*K3 + pSPE2807	0	TraMΔ8 (ΔPm1 and Pm2)

Figure 4.4

TraM and TraMΔ8 are eluted from a gel exclusion column in the same fractions. Western blot analysis was performed of fractions collected from gel exclusion chromatography of crude extracts containing TraM (A) and TraMΔ8 (B). Samples (20μl) were taken from every third fraction and separated by SDS-PAGE. TraM was detected using polyclonal TraM antisera. The solid arrows denote TraM and TraMΔ8 monomers, while the dotted arrows denote TraM and TraMΔ8 dimers.

A.


Fraction number (TraM)
30 32 35 38 40 42 45 48 51 54 57 60 63 66 70 74 112 113

B.

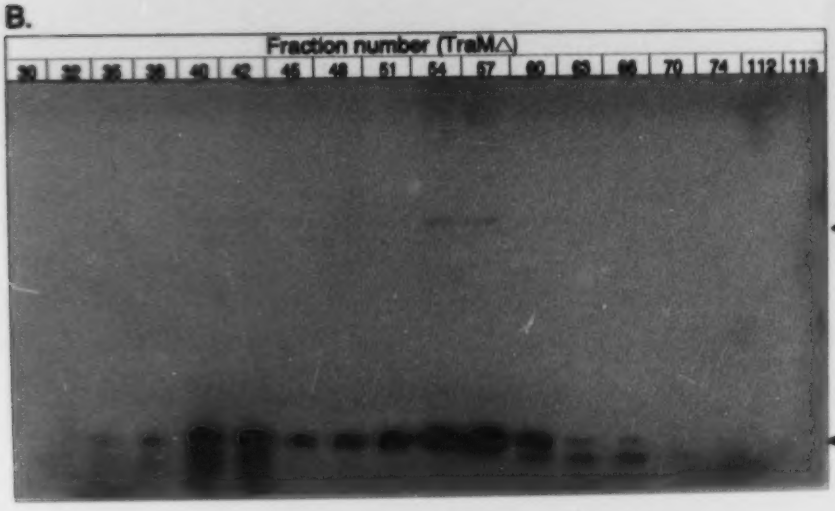
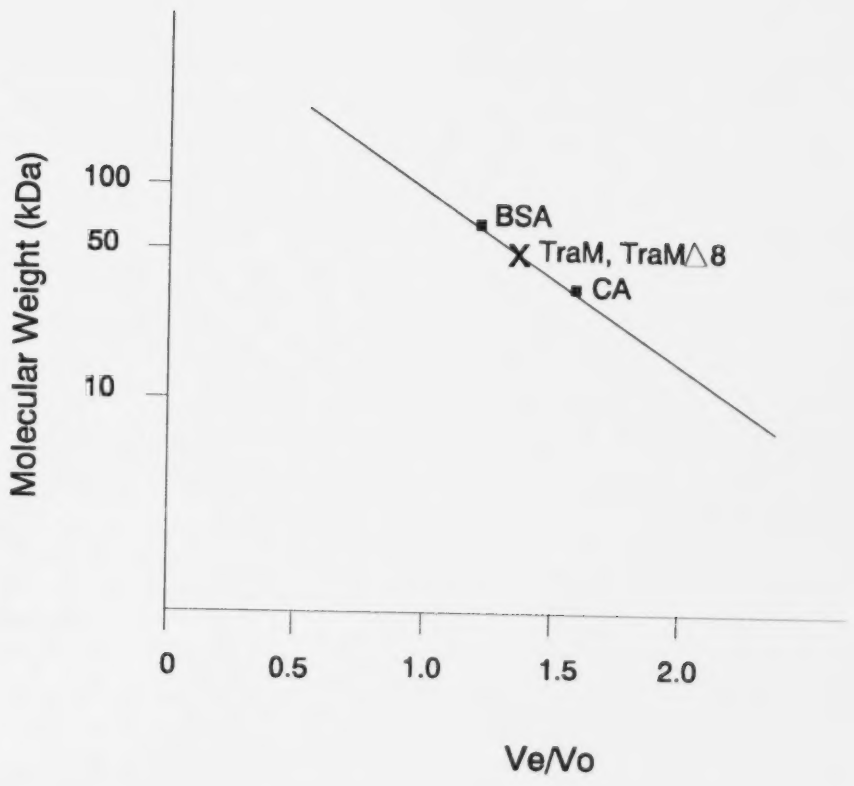

Fraction number (TraMΔ)
30 32 35 38 40 42 45 48 51 54 57 60 63 66 70 74 112 113

Figure 4.5

Calculation of the molecular weight of native TraM and TraMΔ8. Plot of molecular weight (kDa) versus Ve/Vo, where Ve is the elution volume of the protein and Vo is the elution volume of Blue Dextran. Standards used were Bovine Serum Albumin (BSA), MW 66 kDa, and Carbonic Anhydrase (CA), MW 29 kDa.


Molecular Weight (kDa)
100
50
10
BSA
TraM, TraMΔ8
CA
0
0.5
1.0
1.5
2.0
Ve/Vo

removal of the C-terminus of the protein did not affect the quaternary structure, and that TraMΔ8 was still able to form tetramers. No TraM or TraMΔ8 was detected in fractions predicted to contain TraM dimers or monomers, suggesting that the tetramer conformation of these proteins was stable.

V. Gel retardation analysis

The ability of the wild type and mutant proteins to bind to *oriT* DNA was assessed by measuring the electrophoretic mobility of TraM-DNA complexes. The DNA fragments used contained either two (sbmAB, 280 bp) or three (sbmABC, 405 bp) TraM binding sites and were labelled at both ends with the Klenow fragment of DNA polymerase I (refer to Materials and Methods). The protein was allowed to bind to the DNA and the complexes were separated on a non-denaturing gel and detected by autoradiography. Two TraM-*oriT* complexes were detected when increasing amounts of partially-purified wild type TraM protein were added to fragment sbmAB in the presence of competitor DNA (Figure 4.6). The first complex (I) appeared at a protein:DNA ratio of less than 8ng/5 fmol DNA, while the second complex (II) was detected when twice the amount of protein was added. The addition of 80ng of total protein resulted in all of the DNA forming complex II, as no free DNA or complex I could be detected. In contrast, partially purified TraMΔ8 was only able to form a protein-DNA complex at protein:DNA ratios of 800ng:5 fmol DNA. However, this ratio corresponds to the ratio at which crude extract from cells containing the vector alone formed a complex. Since no TraM was present in these extracts, this suggests that this complex is either due to non-specific protein-DNA interactions, or possibly the interaction of the host-encoded protein, IHF, with *oriT* DNA. Similar results were obtained when a fragment containing all three TraM binding sites (sbmABC) was used. Only extracts containing wild type TraM were able to form specific protein:DNA complexes with this fragment (data not shown).

VI. Primer extension analysis of the autoregulatory capacity of the TraMΔ8

Using RNA isolated from cells containing plasmids pOX*traM*K3 and one of pLFR28, pSPE2809, pSPE2309, pSPE2807 or pSPE2307, the ability of TraMΔ8 expressed *in trans* to repress *traM* expression from either of the two *traM* promoters was investigated. Figure 4.7. shows that no *traM* expression from pOX*traM*K3 was detected when *traM* was expressed *in trans* from a multicopy plasmid lacking any *traM* promoters (pSPE2809). A comparison with cells with cells expressing *traM* from a plasmid containing the *traM* promoters (pSPE2309) showed low levels of *traM* expression. This suggested that these transcripts were the result of expression from promoters on the

Figure 4.6

TraMΔ8 fails to bind *oriT* DNA. Gel retardation analysis using TraM and TraMΔ8 was performed. Fragment sbmAB (5 fmol), containing two TraM binding sites, was end-labelled and incubated in the presence of increasing amounts of crude extracts from cells containing no TraM (pT7.4), wild type TraM (pSPE2309-7) or TraMΔ8 (pSPE2307-7). The mobility of the free fragment is represented by the fastest migrating band in the lane containing 0 ng of added protein. Complexes I and II that formed in the presence of TraM are indicated. Protein-DNA complexes were separated on a non-denaturing 5% polyacrylamide gel and detected by autoradiography.

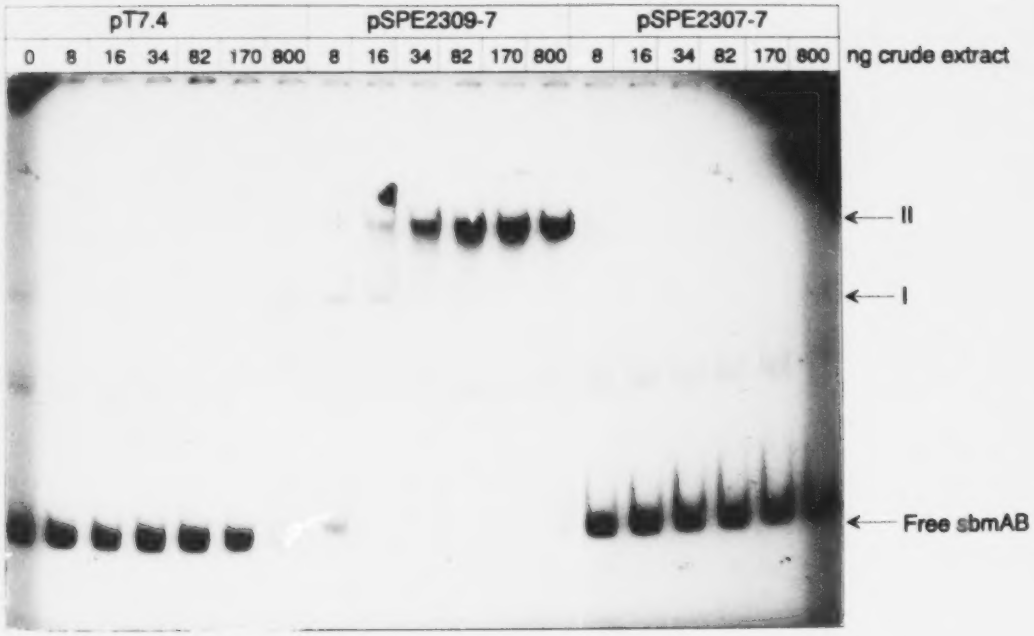
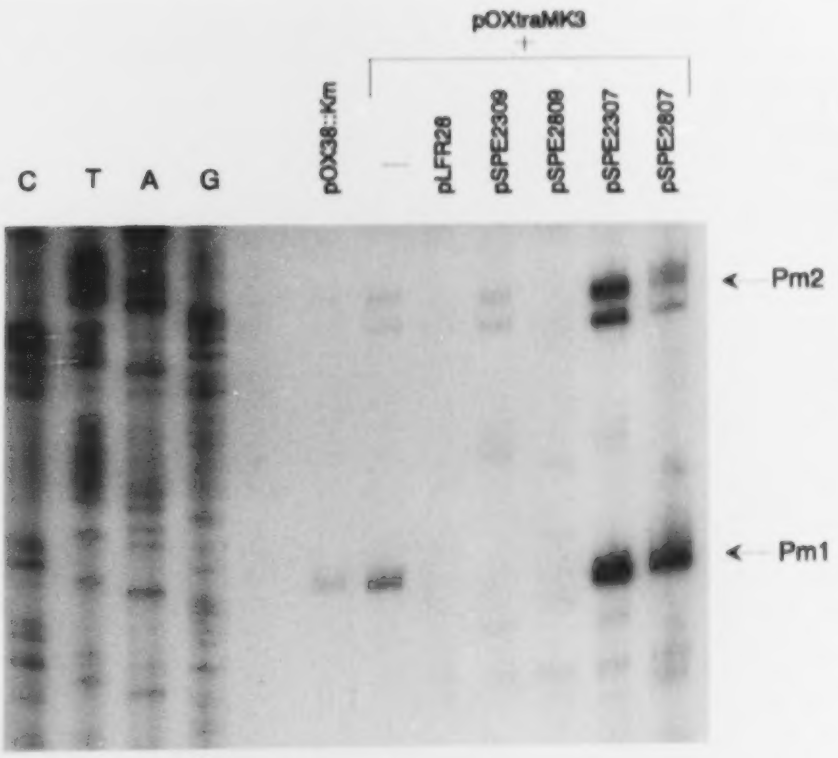

pT7.4
pSPE2309-7
pSPE2307-7
0 8 16 34 82 170 800 8 16 34 82 170 800 8 16 34 82 170 800
ng crude extract
II
I
Free sbmAB

Figure 4.7

TraMΔ8 is unable to repress *traM* promoters *in vivo*. Primer extension analysis was performed on RNA (10μg) isolated from cells containing pOX*traM*K3 and the plasmids indicated, using end-labelled primer SPE8 (Chapter 3). The corresponding nucleotide sequence is given on the right and *traM* promoters Pm1 and Pm2 are indicated.


pOXtraMK3
C
T
A
G
pOX38::Km
—
pLFR28
pSPE2309
pSPE2809
pSPE2307
pSPE2807
Pm2
Pm1

multicopy plasmid rather than pOX*traM*K3. When *traMΔ8* was expressed *in trans* from plasmids bearing a 3' deletion of *traM* (pSPE2307 and pSPE2807), no repression of expression was observed.

C. Discussion

TraM proteins from a number of *traM* alleles, including F, have been shown to migrate anomalously on SDS-PAGE (Di Laurenzio *et al.*, 1991, 1992; Dempsey and Fee, 1990). F TraM has been shown to have a molecular weight of 14,507, but migrates with an apparent molecular weight of 10,950 (Di Laurenzio *et al.*, 1992). The removal of the terminal 8 amino acids of the protein, which removes a number of charged residues, failed to change this anomalous migration. TraMΔ8 migrated with an apparent molecular weight of 9,950, as compared to its predicted molecular weight of 13,371.

TraMΔ8 was incapable of binding *oriT* DNA as demonstrated by gel electrophoretic mobility shift assays. These results were further supported by primer extension analyses which showed that TraMΔ8 was incapable of repressing expression from either Pm1 or Pm2 of pOX*traM*K3. The increased *traM* expression seen from pOX*traM*K3 promoters observed in the presence of pSP2807, which lacks both of the *traM* promoters, suggests that expression of *traM* in pOX*traM*K3 is still repressed to some degree. It is possible that the expression of the first 21 amino acids of TraM in the insertion mutant is sufficient for the formation of the appropriate structure for DNA binding, and hence some repression of the *traM* promoters occurs. If tetramer formation is required for binding, one could speculate that some productive tetramers would form under these conditions. However, in the presence of excess TraMΔ8 produced from multicopy plasmids, only non-productive tetramers would form (composed of TraM and TraMΔ8), therefore no *traM* repression could occur.

The inability of TraMΔ8 to bind *oriT* DNA was somewhat surprising since it has been demonstrated that the N-terminus of TraM is involved in DNA-binding. (Schwab *et al.*, 1993). Mutations in the N-terminus which interfered with the hydrophobic face of the amphiphilic helix produced a protein which was unable to repress expression from a *traM-lacZ* fusion, implicating this region in the DNA-binding properties of TraM. In contrast to the mutant described in the current work, a C-terminal deletion mutant which lacked the terminal third of the protein, was able to repress *traM-lacZ* expression about half as well as wild type protein. However, direct binding of this protein to *oriT* DNA was not demonstrated.

Since TraMΔ8 was unable to bind *oriT* DNA as judged by both gel retardation assays and primer extension analysis, we expected that plasmids expressing this protein would do so

in an unregulated manner. However, immunoblot analysis showed that pSP2309 and pSP2307 expressed similar amounts of TraM and TraMΔ8, respectively. Both plasmids lack the sequence downstream of the *traM* coding region that has been proposed to be a potential transcription terminator sequence (Thompson and Taylor, 1982). Although this region does contain a sequence of dyad symmetry centred 20 nt upstream of a run of T residues that is characteristic of a rho-independent terminator, it lacks the run of continuous GC bp required for the formation of a termination stem-loop structure. Northern blot analysis suggests that the majority of transcripts terminate within this region, as predicted (Chapter 3). However, S1 mapping will be required to identify the position at which *traM* transcripts terminate. The results presented here suggest that this small hairpin loop could be involved in RNA stability.

Previous work (Di Laurenzio, 1992) which utilized sucrose gradient ultracentrifugation and analytical sedimentation equilibrium studies, indicated that TraM occurs as a homotetramer. It has been suggested that at least a single TraM tetramer binds to the distal TraM binding sites of plasmid R1, and that loading of the TraM monomers is co-operative (Schwab *et al*, 1991). No such evidence exists for F TraM and it is not yet known whether a single tetramer binds to each site in *oriT*, or whether two or more sites are bound by a single tetramer. We investigated whether the mutant TraM produced from the C-terminal mutant was capable of forming tetramers, since this may be a pre-requisite for DNA-binding. Western blot analysis of the wild type mutant proteins frequently showed the appearance of a dimer at an apparent molecular weight of 22 kDa, which did not disaggregate during SDS-PAGE. This TraM dimer therefore appears to be extremely stable, and its formation does not involve the C-terminus. The ability of TraMΔ8 to form tetramers, but nevertheless remain unable to bind DNA suggests that tetramer formation is not the only requirement for binding of TraM to *oriT*. In light of these results, one could agree with a postulation by Schwab *et al.* (1991) that the N-terminus of TraM is not directly involved in DNA-binding, but is required for oligomerization. However, the ability of TraM expressed by N-terminal mutants, which were incapable of binding *oriT* DNA, to form tetramers was not tested.

In summary, a C-terminal deletion mutant of TraM, TraMΔ8, has been constructed which is incapable of binding *oriT* DNA and thus incapable of repressing *traM* expression, despite its ability to form tetramers. These results suggest that the C-terminus of F TraM is important in DNA binding. Attempts to generate additional, less severe C-terminal

deletions to identify which amino acids are required to restore DNA-binding activity, are ongoing.

D. Materials and methods

I. Bacterial strains, plasmids and media

The pOX38::Km insertion mutant, pOX*traM*K3 has been described in Chapter 3. For overexpression of proteins from plasmid pT7.4 derivatives, *E.coli* BL21 (DE3) was used (Studier and Moffat, 1986); otherwise all plasmids were transformed into *E.coli* XK1200, described in Chapter 3 (Table 3.2). Recipients used in mating assays were *E. coli* JC3272. Plasmids pT7.4 were provided by Stan Tabor (Department of Biological Chemistry, Harvard Medical School, Boston, MA). Plasmids pLDLF7 (a pUC18 derivative containing *traM*) and pLDF007 (a pT7.4 derivative used to overexpress TraM) have been described previously (Di Laurenzio *et al.*, 1992). Cells were grown in Luria-Bertani broth (Maniatis *et al.*, 1982).

II. Recombinant DNA techniques and reagents

Restriction enzymes, DNA-modifying enzymes and dNTPs were purchased from Boehringer-Mannheim (Laval, Quebec), and used according to the manufacturer's instructions. Vent DNA polymerase was purchased from New England Biolabs, Beverley, MA. [α^{32}P]-dATP and [γ^{32}P]ATP were purchased from Du Pont, Boston, MA. Poly dI-dC DNA, IPTG, kanamycin, ampicillin and rifampicin were obtained from Sigma Chemical Co., St. Louis, MO. Immunoblot analysis was done using an ECL Western blotting detection system, purchased from Amersham Life Science, Buckinghamshire, England. Dithiothreitol (DTT) was purchased from ICN Biochemical, Cleveland, Ohio.

III. Construction of a C-terminal deletion mutant (*traM*Δ8)

The PCR was used to construct a 3' deletion mutant of *traM* which would result in the production of a TraM protein that lacked the terminal 8 amino acids of the wild type protein (TraMΔ8). The sequence of the 3' primer used in the PCR was 5'-TTATCGTTCCATCTCAGATGAT (SPE7), while the sequence of the 5' primer was 5'-GTTTTGTAGGGGTGGTACTGACTA (LFR23) and would thus amplify a fragment containing IHFB, *sbmB*, *sbmA* , and *traM* to nt 820. A second construct bearing the same 5' end was generated, but this fragment was amplified using a 3' primer of sequence 5'-GAATTCTTATTCATCATCATTTTTTG, and therefore generated a fragment expressing wild type *traM*. Conditions for the PCR were as follows: 50pmol of each primer was used to prime synthesis from 50ng of pNY300 template DNA in the presence of 200μmol dNTPs, 10μl 10x Vent buffer (100mM KCl, 200mM Tris-HCl, pH 8.8,

100mM $(NH_4)_2SO_4$, 20mM Mg SO_4, 1% Triton X-100) and 2 U Vent DNA polymerase. The final volume of the reaction was 100µl and the reaction was overlaid with an equal volume of mineral oil. After an initial denaturation step at 94°C for three minutes, the reaction was cycled 30 times at 94°C for 1 minute, 52°C for 1 minute, and 72°C for 2 minutes. A final extension step involved incubation at 72°C for three minutes. Mineral oil was removed by extraction with an equal volume of chloroform and the reaction products were precipitated with 2.5 volumes of 95% ethanol in the presence of 0.3M NaOAc. The pellet was resuspended in 50µl of distilled water and a 10µl aliquot was electrophoresed on a 1% agarose gel. A 582bp fragment (wild type *traM*) and a 552bp fragment (*traM*Δ8) were gel purified as described in Chapter 2. For ligating the blunt-ended fragments encoding wild type *traM* and *traMΔ8*, the fragments were first treated with 10u polynucleotide kinase (PNK) and 1u Klenow before ligation into *Sma*I digested pUC118 (Vieira and Messing, 1987). Treatment with PNK was carried out at 37°C for 30 minutes in the presence of 0.1M Tris-HCl, pH8.0, 10mM $MgCl_2$, 10mM DTT and 1mM ATP in a final volume of 20µl. One unit of Klenow was added and incubation continued for a further 15 minutes. To this was added 0.2µg of *Sma*I digested, phenol extracted and ethanol precipitated pUC118 DNA, resuspended in 5µl of water. Ligation buffer (20µl) (final concentrations 20mM Tris-HCl, pH7.5, 10mM $MgCl_2$, 10mM DTT, 1mM dNTPs and 2mM ATP) was added followed by 1 U T4 DNA ligase. The ligation was carried out at 12°C overnight. Ligation reactions were phenol extracted and precipitated with ethanol and resuspended in 10µl of water. A 1/5th volume of the ligations was electroporated into electrocompetent *E. coli* DH5α and cells were incubated overnight at 37°C on LB plates containing 50µg/ml ampicillin. Analysis of DNA prepared by small scale alkaline lysis (Birnboim and Doly, 1979) confirmed the presence of the appropriate size inserts contained in the vector, and dideoxy sequencing (Sanger, 1977) was used to confirm the presence of fragments carrying either the wild type *traM* gene (plasmid pSPE2309), or a mutant *traM* which lacked the terminal 8 amino acids of the gene (plasmid pSPE2307).

Two additional plasmids expressing wild type *traM* and *traM*Δ8 were constructed for use in primer extension assays to determine whether the mutant protein could repress *traM* expression. These plasmids, pSPE2809 and pSPE2309 respectively, lacked the *traM* binding sites *sbmB* and *sbmA*, and the *traM* promoters Pm1 and Pm2. The 5' primer used in the generation of these constructs had the sequence 5'-CGAATTCGTCCCTGTTTGCATTATGA (LFR28), while the 3' primers were the same as those used to generate pSPE2309 and pSPE2307 for *traM* and *traM*Δ8, respectively.

The fragment encoding wild type *traM* was digested with *Eco*RI and ligated into 0.2µg *Eco*RI digested pUC118, which had been treated with calf intestinal phosphatase (1u) in a 50µl reaction volume for 1 hour at 37°C following restriction. The DNA was phenol extracted and precipitated in ethanol prior to resuspension in water. The fragment encoding the *traM* deletion was treated with Klenow and PNK as described earlier and ligated into *Sma*I-digested pUC118. The plasmids thus generated were called pSPE2809 (containing the wild type *traM* gene) and pSPE2807 (containing *traM*Δ8). A schematic diagram of the plasmids used in this study is shown in Figure 4.2 and general descriptions of the plasmids are given in Table 4.2.

IV. Overexpression of TraM and TraMΔ8

In order to overexpress TraM and TraMΔ8, the *traM* and *traM*Δ8 genes were ligated into the pT7.4 (Tabor and Richardson, 1985) vector. Wild type and mutant *traM*-containing fragments from pSPE2309 and pSPE2307 were excised by digestion with *Bam*H1 and *Eco*R1, and ligated into similarly digested pT7.4 to generate pSPE2309-7 and pSPE2307-7 respectively. Ligations were electrpoporated into electrocompetent *E. coli* BL21 (DE3) cells and colonies containing plasmids of the correct size, as shown by small-scale DNA analysis (Birnboim and Doly, 1979), were chosen for overexpression. Cultures (5ml) were grown to an OD of 1.2 and plasmid-borne proteins were induced by the addition of 100µl of 0.1M IPTG. Cells were incubated at 37°C for up to two hours and 200µl aliquots were drawn at various intervals. Cells were pelleted by centrifugation and frozen on ice until further processing. After boiling in sample buffer (60mM Tris-HCl, pH 6.8, 1% β-mercaptoethanol, 1% SDS, 10% glycerol, and 0.01% bromophenol blue), the samples were loaded onto a 15% PAG and electrophoresed. Gels were stained with Coomassie Brilliant Blue in order to visualize proteins.

V. *In Vitro* Labelling of overexpressed proteins

5ml of LB containing 25µg/ml ampicillin was innoculated with an overnight culture of BL21 containing either pSPE2309-7 or pSPE2307-7. Cells were grown to mid-log phase and 1ml aliquots were centrifuged to pellet the cells. Pellets were washed in 1ml M9 salts (Maniatis *et al.*, 1982), 5mM $MgSO_4$, 0.2% glucose and centrifuged again. After resuspension in the same buffer, 5µl of ampicillin (5mg/ml) was added and the mixture was incubated at 37°C for 1 hour. Expression of plasmid-borne proteins was induced by the addition of 20µl of 0.1M IPTG (final concentration 2mM) and cells were left at 37°C for a further 5 minutes. Further initiation was prevented by the addition of 10µl of rifampicin (20mg/ml) and incubation was continued for a further 5 minutes prior to the

Table 4.2

General characteristics of plasmids used to study the effect of *traM*Δ8. The table indicates the presence of *traM* promoters Pm1 and Pm2, the presence of TraM binding sites and the vector used for each construct.

Plasmid	*traM* Promoters	TraM binding sites	Vector	TraM/TraMΔ8
pLDLF7	Pm1, Pm2	sbmABC	pUC18	TraM
pLDLF007	Pm1, Pm2	sbmABC	pT7.4	TraM
pSPE2309	Pm1, Pm2	sbmAB	pUC118	TraM
pSPE2309-7	Pm1, Pm2	sbmAB	pT7.4	TraM
pSPE2307	Pm1, Pm2	sbmAB	pUC118	TraMΔ8
pSPE2307-7	Pm1, Pm2	sbmAB	pT7.4	TraMΔ8
pSPE2809	None	None	pUC118	TraM
pSPE2807	None	None	pUC118	TraMΔ8

addition of 1µl of ^{35}S-methionine. Aliquots (200µl) were taken at intervals, cells were pelleted by centrifugation and pellets were frozen on ice until further processing. For SDS-PAGE analysis, cells were boiled in sample buffer prior to loading an a 15% PAG. Gels were dried and autoradiography was performed overnight at -70°C using Kodak XAR film.

VI. Mating assays

The ability of the TraMΔ8 protein expressed by plasmids pSPE2807 or pSPE2307 to complement the *traM* insertion mutant, pOX*traM*K3, in mating assays, was assessed. Mating assays were performed as described in Chapter 2, and the mating efficiencies compared to those obtained when *traM* was expressed in *trans* from plasmids pSPE2809 or pSPE2309.

VII. Detection of proteins by immunoblot analysis

Following electrophoresis, proteins were transferred to Immobillon nylon membranes (Millipore) using a Trans-blot Semi-Dry Transfer Cell (Biorad). Conditions for transfer were those recommended by the supplier (12V) and transfer was allowed to proceed for 30 minutes (mini-gels) or 45 minutes (large gels). Blocking reagent used to mimimize non-specific cross-reactivity was 10% skim milk powder (Difco) and the blocking reaction was carried out for 1 hour at room temperature. Immunological detection was done using polyclonal antisera raised against purified TraM and an ECL Detection Kit (Amersham), used according to the manufacturer's instructions.

VIII. Preparation of crude protein extracts for FPLC purification

Cultures (BL21 containing plasmids pSPE2309-7 or pSPE2307-7) were incubated with shaking at 37°C overnight. Plasmid protein production was induced by the addition of IPTG at a final concentration of 1mM and incubation for a further 2 hours. Cells were cooled on ice and centrifuged at 10,000g for 10 minutes. The pellet was washed in 20ml of 20% sucrose, 30mM Tris-HCl, pH8.0 and centrifuged again, before being resuspended in 2.4ml of the same buffer. Lysozyme (16µl of a 12mg/ml solution in 0.25M EDTA) was added (final concentrations 80µg/ml and 1.7mM respectively) and the suspension was incubated on ice for 15 minutes. Two volumes of ice cold distilled water were added and incubation on ice was continued for 30 minutes. Cells were disrupted by three passages through a French Pressure Cell (American Instrument Co., Silver Spring, Md) at 986Kg/cm^2. Cell debris was pelleted by centrifugation at 100,000g for 1.5 hours in an ultracentrifuge (Beckman model L8-M, SW41 rotor, Beckman Corp., Sunnyvale, Ca.).

The supernatant, containing cytoplasmic proteins, was dialysed overnight against TED buffer (50mM Tris-HCl, pH 7.5, 0.1mM EDTA, 1mM DTT), and concentrated in an Amicon concentrator using YM10 membranes which have a molecular weight cut-off of 10kDa (Filtron Technology Corp., Clinton, MA). Protein concentrations were determined using Biorad Protein Detection Reagent.

IX. FPLC purification of TraM and TraMΔ8 from crude protein extracts

Crude protein preparations (100µl) were diluted in 0.2M NaCl to disrupt aggregates prior to loading a 200µl volume on an FPLC gel exclusion column (Highload 16/60 Sephacryl S-200 HR, Pharmacia, Uppsala, Sweden). Proteins were eluted off the column at a rate of 1ml/min in TED buffer, pH 9.0, containing 0.2mM NaCl. Aliquots (20µl) were taken from every third fraction, freeze dried and resuspended in 1x sample buffer prior to electrophoresis on a 15% PAG. Immunological detection was used to identify those fractions that contained TraM. Fractions containing wild type TraM were pooled, as were those shown to contain TraMΔ8. Protein concentrations of the pooled samples were determined using Biorad Detection Reagent, and these pools were used for gel shift analyses.

X. Gel retardation assays

Protein-DNA complexes were detected by the method of Garner and Rezvin (1981). Fragment sbmAB was generated by digesting pLFR23 (Chapter 3) with *Sal*I to obtain a fragment of 280 bp, containing TraM-binding sites *sbmA* and *sbmB*. Fragment sbmABC was obtained by digesting pLFR24 with *Sal*I to obtain a 405 bp fragment containing TraM-binding sites *sbmA*, *sbmB* and *sbmC*. The fragments were purified from an agarose gel as described in Chapter 2, and resuspended in distilled water. DNA was labelled using [α-^{32}P]-dATP and Klenow enzyme in a "3' fill-in" reaction (Maniatis, *et al.*, 1982). The reaction was terminated by heating at 65^{o}C for 10 minutes, proteins were removed by phenol extraction and the DNA precipitated with ethanol. The pellet was resuspended in 100µl distilled water and unincorporated nucleotides were removed by chromatography through a Biogel P-30 (Bio-rad Laboratories Ltd. Mississauga, ON) spin column. Approximately 5 fmol of DNA was used in gel shift assays. For the assays, FPLC purified protein preparations were diluted serially in 1x Retardation Buffer (50mM Tris-HCl, pH 7.5, 10% glycerol, 1mM EDTA, 1mM DTT, 30µg/ml BSA), and increasing amounts of protein were added to the labelled fragment in the presence of 1µg competitor DNA (poly [dI-dC] poly [dI-dC]). As a control, crude protein preparations from cells containing the pT7.4 vector alone were used to detect non-specific binding.

Reactions (20µl total volume in 1x Retardation Buffer) were incubated at 37°C for 10 minutes prior to loading onto a pre-run 5% non-denaturing polyacrylamide gel. Loading took place with the gel running at 300V, whereafter the voltage was decreased to 200V and electrophoresis continued for 2 hours at room temperature. Protein-DNA complexes were detected by exposing the dried gels to Kodak XAR film overnight.

XI. Primer extension analysis

The ability of TraMΔ8 protein to repress expression from either of the two *traM* promoters was measured by doing primer extension anlaysis on RNA extracted from cells containing pOX*traM*K3 and one of pSPE2307 or pSPE2807. Primer extension analysis was performed as described in Chapter 3 and dried gels were autoradiographed overnight at -70°C.

E. References

Birmboim, H. C., and J. Doly. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7:** 1513-1523.

Dempsey, W. B. and B. E. Fee. (199). Integration host factor affects expession of two genes at the conjugal transfer origin of plasmid R100. *Mol. Microbiol.* **4:** 1019-1028.

Di Laurenzio, L. (1992) Protein-DNA interactions at *oriT* regions. Ph.D. Thesis. University of Alberta, Edmonton, Canada.

Di Laurenzio, L., B. B. Finlay, L. S. Frost, and W. Paranchych. (1991) Characterization of the oriT region of the IncFV plasmid pED208. *Mol. Microbiol.* **5:** 1779-1790.

Di Laurenzio, L., L. S. Frost. and W. Paranchych. (1992) The TraM protein of the conjugative plasmid F binds to the origin of transfer of the F and ColE1 plasmids. *Mol. Microbiol.* **6:** 2951-2959.

Garner, M. M., and A. Revzin. (1981) Gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Res.* **9:** 3047: 3060.

Göldner, A., H. Graus, and G. Högenauer. (1987) The origin of transfer of P307. *Plasmid* **18:** 76-83.

Koronakis, V. E., E. Bauer, and G. Högenauer. (1985) The *traM* gene of the resistance plasmid R1: comparison with the corresponding sequence of the *Escherichia coli* F factor. *Gene* **36:** 79-86.

Maniatis, T., E. F. Fritsch, and J. Sambrook. (1982) Molecular Cloning. A Laboratory Manual, (1st ed.) Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Sambrook, J., E. F. Fritsch, and T. Maniatis. (1989) Molecular Cloning. A Laboratory Manual, (2nd ed.) Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Sanger, F., S. Nicklen, and A. R. Coulsen. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74:** 5463-5467.

Schwab, M., R. Fratte, B. Jauk, G. Koraimann, and G. Högenauer. (1994) Functional analysis of protein TraM from the resistance plasmid R1. (Abstract). Plasmid Biology. Banff, Alberta, Canada

Schwab, M., H. Gruber, and G. Högenauer. (1991) The TraM protein of plasmid R1 is a DNA-binding protein. *Mol. Microbiol.* **5:** 493-446.

Schwab, M., H. Reisezein, and G. Högenauer. (1993) TraM of plasmid R1 regulates its own expression. *Mol. Microbiol.* **7:** 795-803.

Studier, F. W., and B. A. Moffat. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189:** 113-130.

Tabor, S., and C. C. Richardson. (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* **82:** 1074-1078.

Thompson, R., and L. Taylor. (1982) Promoter mapping and DNA sequencing of the F plasmid transfer genes *traM* and *traJ*. *Mol. Gen. Genet.* **188:** 513-518.

Vieira, J., and J. Messing. (1987) Production of single-stranded plasmid DNA. *Meth. Enzymol.* **153:** 3-11.

Willetts. N. (1981) Sites and systems for conjugal DNA transfer in bacteria. *In* Molecular Biology, Pathogenicity, and Ecology of Bacterial Plasmids Levy, p. 207-215. S. B., Clowes, R. C., and Koening, E.L. (eds). New York: Plenum Press.

Willetts, N., and J. Maule. (1986) Specificities of IncF conjugation genes. *Genet. Res.* **47:** 1-11.

Chapter 5

The binding of TraM to *oriT* DNA

A. Introduction

The TraM proteins of plasmids R1, R100, pED208 and F have all been shown to bind *oriT* DNA (Schwab *et al.*, 1991; Abo *et al.*, 1991; Di Laurenzio *et al.*, 1991, 1992). The nature of the interaction of TraM with DNA appears to be complex however, with unusual binding characteristics demonstrated for R1, pED208 and F on gel retardation assays. Although DNaseI footprinting experiments demonstrated two regions of R1 *oriT* that were protected by TraM, gel shift assays showed only one dramatic shift when pure protein was allowed to bind to a fragment containing both TraM binding sites (Schwab *et al.*, 1991). Only when a fragment containing a single TraM binding site was used, were several steps of retardation that were dependent on the protein concentration observed. These steps were interpreted as the successive loading of TraM monomers, possibly in a co-operative manner, to the single site.

In contrast, gel shift assays on the *oriT* region of pED208 using purified TraM, showed a broad band of decreasing mobility as the protein concentration was increased (Di Laurenzio *et al.*, 1991). No discrete steps representing the ordered loading of TraM sites were observed, and DNA in both the bound and unbound states was never visualized at a given protein concentration. These results were interpreted to mean that binding of pED208 TraM to *oriT* was not co-operative and that DNA-protein complexes underwent continuous exchange between the three binding sites, leading to the diffuse bands that were observed.

Binding of F TraM to *oriT* DNA resulted in the formation of 4 complexes with altered mobility on gel retardation assays (Di Laurenzio *et al.*, 1992). The fourth complex (IV) however, was detected only at high protein:DNA ratios (60,000). Since only 3 protected regions of *oriT* DNA were detected in DNaseI footprinting experiments of the region using crude cell extracts containing TraM, it was concluded that complex IV represented the formation of protein aggregates on already complexed DNA. Like the pED208 system, DNA in both the bound and the unbound states was never detected. These results, together with evidence that TraM aggregates readily (Schwab *et al.*, 1991; Di Laurenzio *et al.*, 1992) led to the suggestion that TraM forms nucleosome-like structures, explaining the apparent requirement for large amounts of TraM protein in retardation assays.

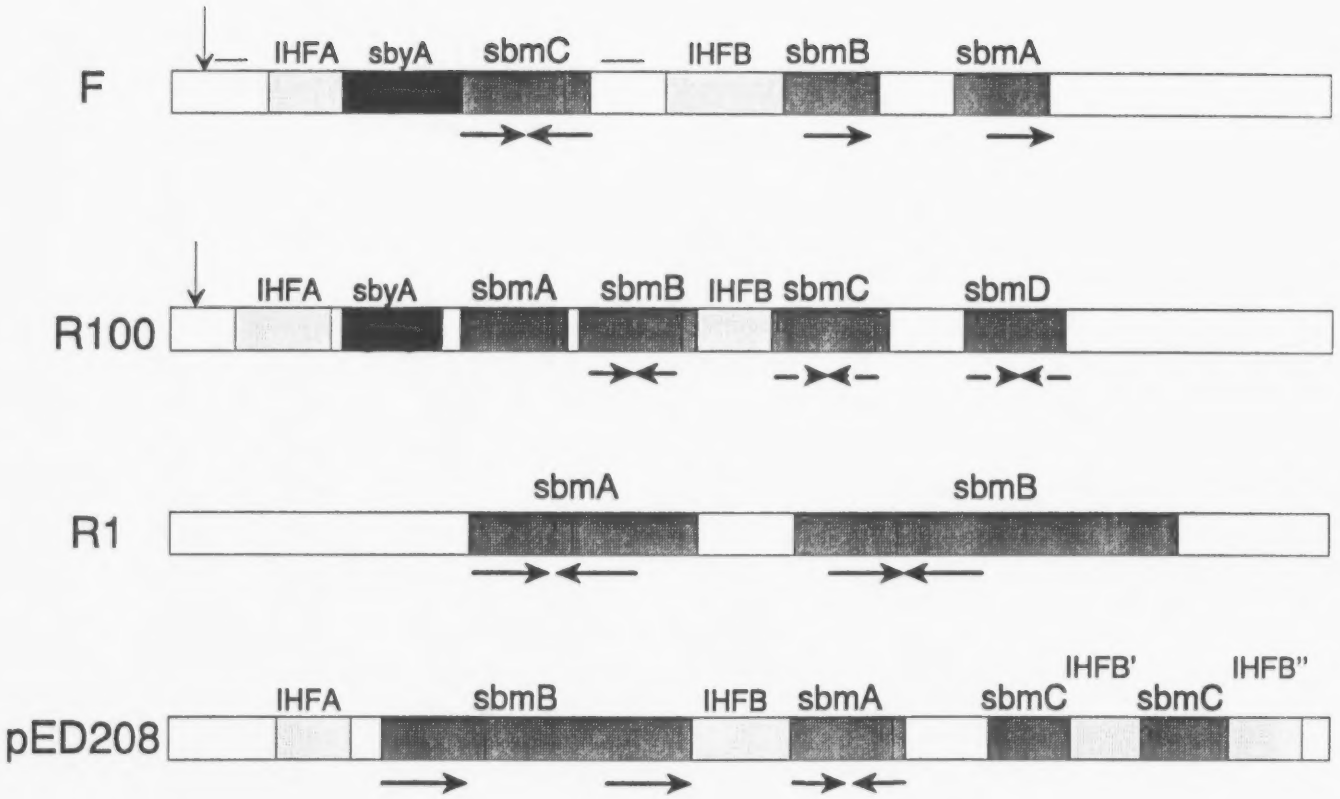
Taken together, the above results all suggest a complex interaction of multiple molecules of TraM with *oriT* DNA. In addition, an examination of the binding sites for TraM from the various plasmids, suggests a complex recognition sequence for this protein. A summary of TraM-*oriT* interactions for the four IncF plasmids F, pED208, R100 and R1 are shown in Figure 5.1. pED208 TraM binds to both an inverted repeat and two direct repeats within the *oriT* region. These binding sites all contain multiple *Hinf*I-like recognition sequences (GANTC) that are spaced at regular intervals 11-12 bp apart. Similarly, plasmid R1 contains a series of *Hinf*I-like recognition sites spaced at regular intervals within the TraM binding sites Both R1 and R100 TraM bind to inverted repeats in their respective *oriT* regions. R100 *oriT*, like F *oriT* lacks any *Hinf*I-like sequences. However, F TraM also binds to an inverted repeat (*sbmC*) and to two direct repeats (*sbmA* and *sbmB*). A recognition consensus sequence of GGPyGC has been proposed for the F TraM and this sequence appears 10 times within the three binding sites. An unusual property of F TraM binding is its apparently different affinities for the upper and lower strands of *oriT* DNA. DNaseI footprinting experiments showed that the lower strand was protected only in the -35 region of Pm1 and to a lesser extent in the -10 region. In contrast, the upper strand was protected over both the -10 and -35 regions. A noticeable feature of the F TraM binding sites is the high proportion of GC residues in comparison with the rest of the *oriT* region, which is extremely AT-rich.

All gel retardation assays on F TraM were done using a fragment of *oriT* which contained all three TraM binding sites. The apparent affinity of TraM for *sbmA*, *sbmB*, and *sbmC* was inferred from the results of DNaseI footprinting experiments, which suggest that increasing concentrations of protein were required for TraM to bind to *sbmA*, *sbmB*, and *sbmC*, respectively (Di Laurenzio *et al.*, 1992).

In order to further investigate the interactions between TraM and *oriT*, the three TraM binding sites were isolated and the binding of TraM to fragments containing all three, only two, or only one of the three sites was investigated. In addition, the role of TraY, which also binds within the *oriT* region of F, in the *in vitro* binding of TraM to *oriT*, was investigated. Protein-DNA complexes formed with the pure protein were compared with those formed with TraM-containing crude extracts, which more clearly resemble the *in vivo* situation.

Figure 5.1

A comparison of the TraM binding sites to *oriT* in four *traM* alleles, F, R100, R1 and pED208. Direct and inverted repeats are indicated by bold arrows, and TraM binding sites are shaded. IHF binding sites are indicated (stippled boxes) as well as TraY binding sites, where known (solid). The sites of nicking in F and R100 are given by arrows, and the two overlines in F *oriT* represent IB1 (internal bend) and IB2.


F
IHFA
sbyA
sbmC
IHFB
sbmB
sbmA
R100
IHFA
sbyA
sbmA
sbmB
IHFB
sbmC
sbmD
R1
sbmA
sbmB
pED208
IHFA
sbmB
IHFB
sbmA
sbmC
IHFB'
sbmC
IHFB''

B. Results

I. The binding of pure TraM to fragment sbmABC

Fragment sbmABC was generated by digesting pLFR24 with *Sal*I. The 405 bp fragment was purified and labelled at both ends in a 3' "fill in" reaction using the Klenow fragment of DNA polymerase I. Five fmol of the fragment was incubated with increasing amounts of pure TraM protein and the complexes were separated on a non-denaturing polyacrylamide gel. A consistent feature of fragments isolated from the *oriT* region was the appearance of a second band, with lower mobility than the free fragment (Figure 5.2). This band was resistant to reisolation and heating and was interpreted as a portion of the free fragment which had taken up an alternative conformation to the bulk of the sample when released from the constraints of the supercoiled plasmid. A variety of means were used in an attempt to stabilise this fragment in its original conformation, but even the addition of 0.3M NaCl failed to have any effect. Unlike the free fragment, this fragment only bound TraM at high DNA:protein ratios (molar ratio 1:4 x 10^3), suggesting a non-specific interaction. Any TraM binding sites present in this fragment must be hidden or not accessible in this conformation.

The free fragment sbmABC formed a complex (I) with pure TraM at a molar ratio of DNA:protein of 1:4 (Figure 5.2), corresponding to the proposed tetrameric quaternary structure of the protein. In contrast to a previous report (Di Laurenzio *et al.*, 1992), DNA was detected in both the bound and unbound states. The addition of twice as much protein (1.2ng/ 40fmol) resulted in the formation of a second complex (II), at a molar ratio of 1:8. Further increases in TraM did not yield additional specific complexes until the DNA:protein ratio was increased above 1:400, where small incremental steps of decreasing mobility were observed. This was interpreted as the product of protein aggregation onto already complexed DNA, rather than specific DNA:protein interactions. These results suggest that a single tetramer binds to each of two sites within fragment sbmABC. Based on the affinity of TraM for each of the three proposed binding sites in *oriT*, we conclude that complex I represents a tetramer of TraM binding to *sbmA*, while complex II represents a second tetramer of TraM binding to *sbmB*.

II. The binding of TraM in crude extracts to *oriT* DNA

Using crude extracts containing TraM and the same DNA fragment, only a single band of reduced mobility was observed (Figure 5.3). Unlike the results observed with pure protein, no DNA in both the bound and unbound states in the same sample was detected.

Figure 5.2

Binding of pure TraM to fragment sbmABC. The free fragment (5 fmol) was incubated with increasing amounts of total protein and protein-DNA complexes were separated on a 5% non-denaturing SDS polyacrylamide gel. The mobility of the free fragment is indicated, and TraM-*oriT* complexes I and II are marked.

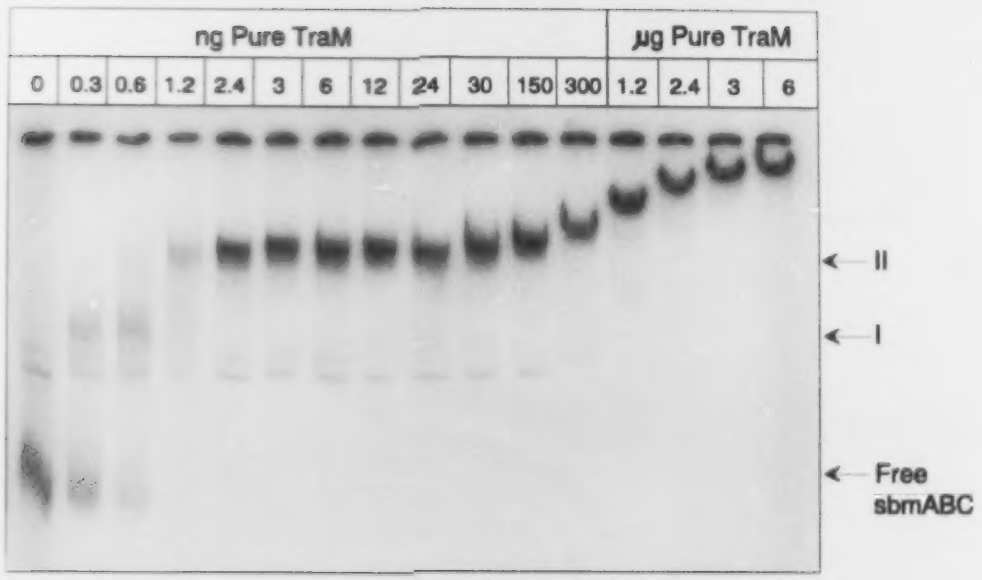
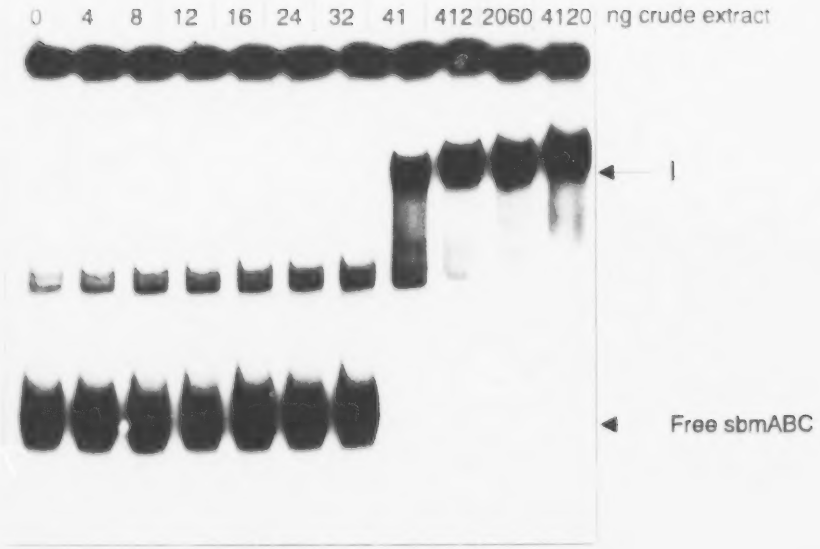

ng Pure TraM
µg Pure TraM
0
0.3
0.6
1.2
2.4
3
6
12
24
30
150
300
1.2
2.4
3
6
II
I
Free
sbmABC

Figure 5.3

Binding of TraM in crude protein extracts to fragment sbmABC. The free fragment (5 fmol) was incubated with increasing amounts of total protein and protein-DNA complexes were separated on a 5% non-denaturing SDS polyacrylamide gel. The DNA:protein complex is indicated (I).


0 4 8 12 16 24 32 41 412 2060 4120 ng crude extract
I
Free sbmABC

The mobility shift was remarkably sudden, as no complexes were detected after the addition of 32ng of protein, but all DNA appeared to be complexed with protein after the addition of 42ng of protein. This DNA-protein complex corresponds in mobility to complex II formed with pure protein. The presence of additional cellular proteins then, appears to cause co-operative binding of TraM to *oriT* DNA.

A fragment containing only two of the three TraM binding sites, fragment sbmBC, was retarded by TraM-containing crude extracts (Figure 5.4). Although a portion of the DNA population was complexed with protein by the addition of 41ng of crude extract, 412 ng of protein was required to complex the entire population. This contrasts to the amount of protein required to form a complex with the same molar quantity of fragment sbmABC, which required only 41ng of crude extract to complex the entire population of DNA.

Fragment sbmC contains only one TraM binding site, previously identified by DNaseI footprinting experiments (Di Laurenzio *et al.*, 1992). This site had low affinity for TraM, as it was protected from cleavage by DNaseI only at high protein concentrations. Similarly, Figure 5.4 shows that this fragment forms a weak complex with TraM-containing crude extracts only at high protein concentrations (molar ratio of DNA:total protein of approximately 1:1400). This complex migrates as a smear, suggesting that the protein-DNA interactions are continuously forming and dissociating.

III. The effect of TraY on TraM-*oriT* interactions

The low affinity of *sbmC* for TraM suggests that either this site is not physiologically relevant, or conditions used in gel shift assays were not optimal for the formation of this complex. The proximity of the TraY binding site, *sbyA*, to *sbmC*, suggested that perhaps the binding of TraY to *oriT* would influence TraM binding to *oriT*. This was investigated by doing gel shift assays with TraM-containing crude extracts, and, in addition, sufficient purified TraY to form a TraY-*oriT* complex. Figure 5.5A shows the effect of pre-incubating fragment sbmBC with TraY, prior to the addition of increasing amounts of TraM-containing crude extract. An initial complex migrating with decreased mobility compared to the free fragment was the result of TraY-*oriT* complex formation. A second complex which migrated with an even slower mobility was formed in the presence of 41ng of crude extract. A comparison with the amount of crude extract required to complex all of fragment sbmBC in the absence of TraY, shows that slightly less protein is required to form the TraM-*oriT* complex in the presence of TraY.

Figure 5.4

Binding of TraM in crude protein extracts to fragment sbmBC and fragment sbmC. The free fragments (5 fmol of each), which migrate with slightly different mobilities as a result of their slightly different lengths, were incubated with increasing amounts of total protein and protein-DNA complexes were separated on a 5% non-denaturing SDS polyacrylamide gel. The protein:DNA complex formed with sbmBC is indicated (I) and the unstable complexes formed with fragment sbmC are bracketed.

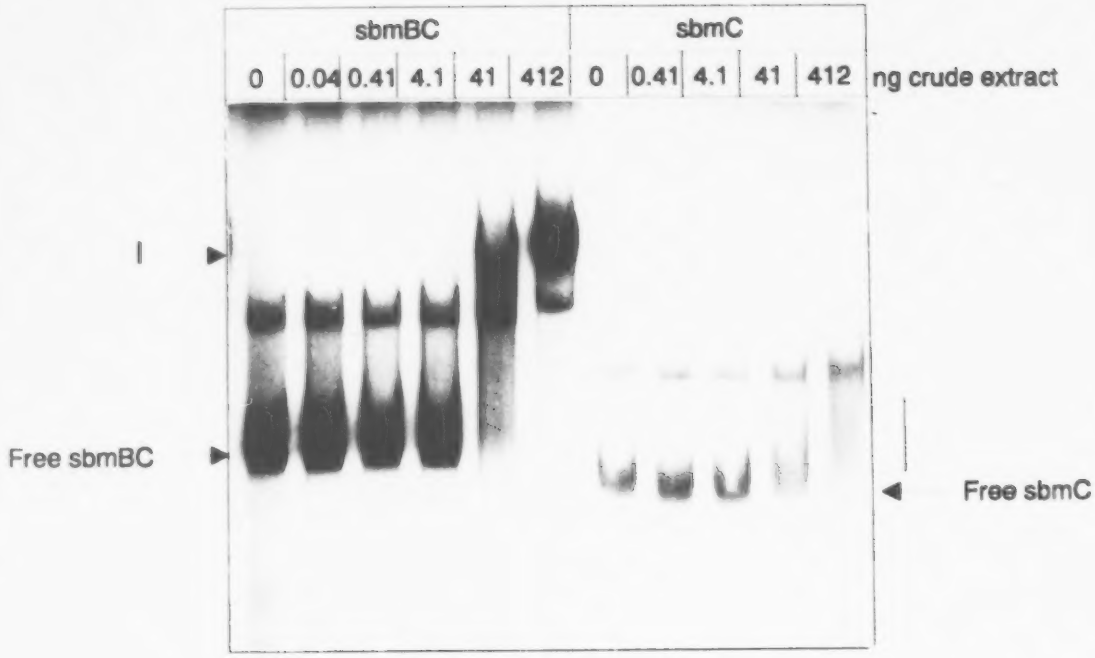
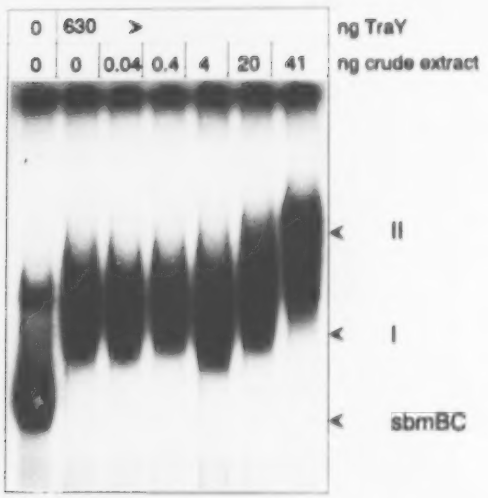

sbmBC
sbmC
0
0.04
0.41
4.1
41
412
0
0.41
4.1
41
412
ng crude extract
I
Free sbmBC
Free sbmC

Figure 5.5

The effect of TraY on the binding of TraM to *oriT* DNA. Five fmol of each of fragments sbmBC (A) and sbmC (B) were incubated in increasing amounts of TraM-containing crude extracts in the presence of an excess amount of pure TraY. Protein-DNA complexes were separated on a 5% non-denaturing SDS polyacrylamide gel. The TraY:DNA complexes are labelled I and the TraM:TraY:DNA complexes are labelled II.

A.


0
630
>
ng TraY
0
0
0.04
0.4
4
20
41
ng crude extract
II
I
sbmBC

B.

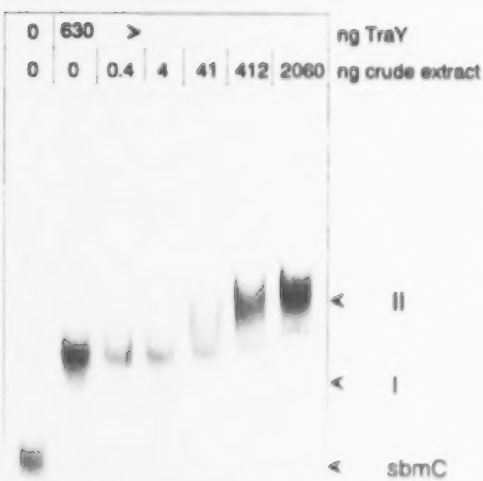

0
630
>
ng TraY
0
0
0.4
4
41
412
2060
ng crude extract
II
I
sbmC

Figure 5.5B shows the effect of pre-incubating fragment sbmC with TraY prior to the addition of increasing amounts of crude extract containing TraM. A TraY-*oriT* complex was detected as a (smeared) band of reduced mobility in the presence of TraY. As the amount of crude extract added was increased above 4ng, the mobility of the fragment was reduced further, until at the addition of 412ng of crude extract, the smear began to sharpen into a band. Although the high concentration of protein required to observe the formation of this sharp band suggests that this may be a non-specific interaction, clearly the presence of TraY enhances the formation of this protein-DNA complex. Since all DNA-binding proteins show non-specific binding to DNA at high concentrations (Lane *et al.*, 1992), this effect could be the result of increasing the concentration of DNA-binding proteins in the reaction by the addition of TraY. Alternatively, TraY may play a role in increasing the affinity of TraM for *sbmC*.

C. Discussion

A careful analysis of the binding of TraM to the *oriT* region has led to a re-evaluation of previous results of gel retardation analyses. By using small incremental increases in the amount of pure TraM added to a constant amount of sbmABC DNA, only 2 complexes resembling specific TraM-*oriT* interactions were detected. These complexes were formed at DNA:protein ratios of 1:4 and 1:8 respectively, in contrast to previous reports which stated that enormous quantities of TraM were required even for initial complex formation (Di Laurenzio *et al.*, 1992). The proposed tetrameric conformation of native TraM is consistent with the results presented here, which suggest that a single tetramer binds to each of the two sites.

The co-operative nature of TraM binding is apparent in the limited conditions under which DNA was observed in both the bound and unbound states. Only when pure TraM, or a fragment containing only two of the three TraM binding sites, was used (see Figures 5.2 and 4.6), was both free and complexed DNA detected. Note that one tenth of the amount of DNA was used in the retardation assay presented in Figure 4.6, therefore less protein was required to form a complex than in Figure 5.3. In addition, using crude extracts containing TraM, only a single dramatic shift was observed with a fragment containing all three binding sites. One can conclude from the results presented here that components present in the crude extract interact with the region between *sbmB* and *oriT* to cause co-operative binding of TraM to *oriT* DNA. Two elements are found in this region which may be responsible for altering DNA structure, and thus affecting TraM binding. The first is the sequence-directed bend IB2, which lies adjacent to *sbmC*, and the second is IHF site B, adjacent to *sbmB*. IHF has been shown to bind to site IHFB and bend DNA (Tsai *et al.*, 1990). Although IHFB is present on fragment sbmAB used in gel retardation assays, its presence on the extreme end of the linear fragment may not have as dramatic an effect on DNA conformation as when it is located at an internal site within the DNA fragment, as in fragment sbmABC. A cautionary word is given here on the interpretation of the absence of bands of intermediate moblity as a co-operative effect of protein binding. It has been demonstrated (Kleinschmidt *et al.*,1991) that weak interactions of the Tet repressor protein with one of two operator sites on a fragment, may be lost in favour of double occupancy on a previously singly-occupied site. So while dissociation of a doubly-occupied complex followed by re-association can result only in the formation of the original complex, dissociation from a singly-occupied complex in the presence of a similar fragment, may yield a doubly-occupied fragment and a free

fragment upon re-association. These anomalies that occur in gel retardation assays should always be borne in mind when interpreting such data.

Under the conditions described here, a TraM-*oriT* complex with fragment sbmC was formed only at high protein concentrations, and this complex did not appear stable, as evidenced by the smeared appearance of the complex. Although the presence of TraY did appear to increase both the rate of formation of the complex as well as the stability of the complex, this interaction still seemed relatively weak in comparison with TraM interactions with *sbmA* and *sbmB*. Clearly, if this complex does form *in vivo*, additional factors not identified in the data presented here may be required to stabilise the complex. In this regard, considerably more TraY protein was required to form a complex than previously reported (Lahue and Matson, 1990), suggesting that the protein may not have been fully active. The absolute requirement of *sbmC* for transfer (Fu *et al.*, 1991), suggests that TraM interaction with this site may provide the link between the proposed signalling function of TraM (Willetts and Wilkins, 1984), and the initiation of transfer. The identification of the requirements for binding of TraM to *sbmC* may provide the clue to how TraM signals that transfer may begin.

With a more thorough understanding of the nature of TraM-*oriT* interactions, a number of questions still remain unanswered. A major consideration is that all gel retardation assays reported thus far have been done using linear fragments of DNA. It is quite possible that the presence of a supercoiled template would alter the binding characteristics considerably.

D. Materials and methods

I. Pure protein preparation

F TraM was prepared by L. Di Laurenzio, as previously described (Di Laurenzio *et al.*, 1992). Briefly, TraM was overexpressed, cells were lysed by treatment with lysozyme and passage through a French pressure cell. Stepwise ammonium sulfate precipitation of proteins in the cell lysate was used to separate the bulk of the cellular proteins from TraM. Passage over an anion exchange column followed by a Blue Sepharose (Biorad) column resulted in the identification of fractions containing pure TraM protein. Purified TraY protein was the kind gift of S. Matson.

II. Preparation of crude extracts containing TraM

Crude extract preparation was described in Chapter 4. These extracts were semi-purifed over a gel exclusion column, and fractions containing proteins ranging in molecular weight from 28 kDa to 110 kDa were pooled for use in gel retardation assays. The pooled proteins were stored at 4°C in the presence of 0.5mM phenylmethylsulphonyl flouride (Sigma Chemical Co. St. Louis, MO).

III. Preparation of fragments used in gel retardation assays

Fragment sbmABC was prepared as described in Chapter 4. Fragment sbmBC was prepared by digesting pNY300 with *Xba*I and *Hae*II, to generate a 385 bp fragment containing both *sbmB* and *sbmC*. To prepare a fragment containing only *sbmC*, pNY300 was digested with *Sal*I to yield a 542 bp fragment containing all three TraM binding sites. This was subsequently digested with *Rsa*I and the resulting 295 bp fragment containing *sbmC* was isolated. All fragments were separated on an 8% polyacrylamide gel. The gel slice containing the fragment of interest was cut out and the DNA was purified by electroelution (Gobel *et al.*, 1987) using an Elutrap electro-separation system (Schleicher and Schuell, Keene, NH). The DNA was extracted with phenol, precipitated in ethanol in the presence of glycogen and resuspended in double distilled water or TE (10mM Tris-HCl, pH 8.0, 1mM EDTA). All fragments were labelled in a 3' "fill in" reaction (Maniatis *et al.*, 1982) using the Klenow fragment of DNA polymerase I and [α ^{32}P]-dATP. Unlabelled fragments were quantitated by comparison of ethidium bromide flourescence to that of known quantities of DNA of similar size.

IV. Gel retardation assays

DNA-protein interactions were assayed as described in Chapter 4. All DNA-protein incubations were done in the presence of an excess (1μg per reaction) amount of poly [dI-dC] poly [dI-dC] DNA (Sigma Chemical Co., St. Louis, MO) as a competitor to decrease non-specific interactions. Initially, buffer used for retardation assays was 25mM HEPES, pH7.8, 50mM KCl, 0.5mM DTT, 0.05mM EDTA, 5% glycerol and 0.5mM PMSF. DNA was incubated in the presence of protein for 30 minutes at room temperature, prior to separation of the complexes on a pre-run 5% polyacrylamide gel. However, since TraY was found not to bind *oriT* DNA under these conditions, the conditions were altered. Complexes formed between TraM and *oriT* DNA in buffer containing 50mM Tris-HCl, pH 7.5, 10% glycerol, 1mM EDTA, 1mM DTT, 30μg/ml BSA at 37°C for 10 minutes, were identical to those formed under the previous conditions. In addition, TraY bound *oriT* DNA under these conditions, so all reactions were subsequently carried out using these conditions. Following electrophoresis, dried gels were subjected to autoradiography.

E. References

Abo, T., S. Inamoto, and E. Ohtsubo. (1991) Specific DNA binding of the TraM protein to the *oriT* region of plasmid R100. *J. Bacteriol.* **173:** 6347-6354.

Di Laurenzio, L., B. B. Finlay, L. S. Frost, and W. Paranchych. (1991) Characterization of the *oriT* region of the IncFV plasmid pED208. *Mol. Microbiol.* **5:** 1779-1790.

Di Laurenzio, L., L. S. Frost, and W. Paranchych. (1992) The TraM protein of the conjugative plasmid F binds to the origin of transfer of the F and ColE1 plasmids. *Mol. Microbiol.* **6:** 2951-2959.

Fu, F. Y.-H.., M.-M. Tsai, Y. Luo, and R. C. Deonier. (1991) Deletion analysis of the F plasmid *oriT* locus. *J. Bacteriol.* **173:** 1012-1020.

Gobel, U., R. Maas, U. Bantel-Schaal, and A. Clad. (1987) Rapid and quantitative electroelution and electrodialysis of DNA from gels. *J. Biochemical and Biophys. Methods.* **14:** 245-260.

Kleinschmidt, C., K. Tovar, and W. Hillen. (1991) Computer simulations and experimental studies of gel mobility patterns for weak and strong non-cooperative protein binding to two targets on the same DNA: application to binding of Tet repressor variants to multiple and single *tet* operator sites. *Nucleic Acids Res.* **19:** 1021-1028.

Lahue, E.E., and S. W. Matson. (1990) Purified *E. coli* F-factor TraY protein binds *oriT*. *J. Bacteriol.* **172:** 1385-1391.

Lane, D., P. Prentki, and M. Chandler. (1992) Use of gel retardation to analyze protein-nucleic acid interactions. *Microbiol. Rev.* **56:** 509-528.

Maniatis, T., E. F. Fritsch, and J. Sambrook. (1982) Molecular Cloning. A Laboratory Manual. (1st ed.) Cold Spring Harbor, New York :Cold Spring Harbor Laboratory Press.

Schwab, M., H. Gruber, and G. Högenauer. (1991) The TraM protein of plasmid R1 is a DNA-binding protein. *Mol. Microbiol.* **5:** 439-446.

Tsai, M.-M, Y. H. F. Fu, and R. C. Deonier. (1990) Intrinsic bends and integration host factor binding at F plasmid *oriT*. *J. Bacteriol.* **172:** 4603-4609.

Willetts, N., and B. Wilkins. (1984) Processing of plasmid DNA during bacterial conjugation. *Microbiol. Rev.* **48:** 24-41.

Chapter 6

General Discussion

A. General Discussion

I. Characteristics of TraM

TraM has been assigned the role of signalling that a stable mating pair has formed and that DNA transfer can begin (Willetts and Wilkins, 1984). This conclusion was drawn on the basis of a study of the characteristics of a *traM* mutant, JCFL102 (Achtman *et al.*, 1982), which demonstrated that cells expressing this plasmid were phage sensitive and therefore produced functional pili, but were unable to transfer DNA. In addition, these cells were able to form stable mating aggregates (Achtman and Skurray, 1977) and *in vitro* nicking assays showed that *traM* was not required for nicking at *oriT* (Everett and Willetts, 1980).

In recent years, much physical evidence has been gathered on the nature of the *traM* gene product from the conjugative plasmids F, pED208 and R1 (Di Laurenzio, *et al.*, 1991, 1992a; Schwab *et al.*, 1991). The protein is known to bind DNA , despite the absence of an obvious helix-turn-helix motif common to many prokaryotic DNA-binding proteins. Secondary structure predictions of F and pED208 TraM showed a protein with a high helical content, with a few of the helical regions separated by short regions of turns or coils (Di Laurenzio, 1992). In addition, the protein was predicted to be flexible and have three main regions exposed on the surface: a basic region at the amino terminus, a charged region in the middle of the polypeptide, and an acidic region at the carboxy terminus. A similar secondary and tertiary structure was predicted for F TraM and pED208 TraM, which shows the least sequence homology with the other members of F-like *traM* alleles. This suggests that all F-like TraM proteins are likely to have the same secondary and quaternary structures and thus bind DNA in a similar fashion.

II. TraM-DNA interactions

Binding of TraM to *oriT* DNA was demonstrated to be co-operative when present in crude extracts lacking other *tra* proteins (Chapter 5), suggesting that a host-encoded factor affects the binding of TraM to *oriT* DNA. Since two IHF-binding sites have been identified within *oriT* of F, (Tsai *et al.*, 1990), this protein is a strong candidate for this role. In particular, IHFB lies between *sbmB* and *sbmC*, and the binding of IHF at this site could be responsible for affecting the binding of TraM to *oriT* DNA. IHF binding sites have also been identified within the *oriT* regions of plasmids pED208 and R100 (Di Laurenzio *et al.*, 1992b; Dempsey and Fee, 1990), emphasizing the important role that IHF plays at *oriT*. Molar ratios suggested that a single TraM tetramer binds each of *sbmA*

and *sbmB* (Chapter 5). The only conditions under which TraM was found to bind *sbmC* was when a high concentration of TraM-containing crude extracts was added to a fragment containing *sbmC*. Presumably, if this site is occupied *in vivo*, other *tra* functions play a role in either derepressing *traM* expression to increase the intracellular concentration of TraM, or altering the structure of *oriT* to increase the efficiency with which TraM binds to *sbmC*. Results presented in Chapter 5 suggest that TraY may play a role in the binding of TraM to *sbmC*. TraY has been shown to bend DNA upon binding (Luo *et al.*, 1994), an effect which may alter the structure of *oriT* to increase the efficency of TraM binding to *sbmC*.

III. *oriT* secondary structure

The *oriT* region contains numerous features that are implicated in determining a specific secondary structure, and which could be required for the control of transfer. As previously mentioned, IHF binds to two sites within F *oriT* DNA, and is known to bend DNA by 140° upon binding (Yang and Nash, 1989). If these two bends were in opposite planes, this would cause the *oriT* region to adopt a Z-like conformation. In addition, two sequence-directed bends have been identified, and together cause a bend of about 50° centred at nt 245, within *sbmC*. Since TraY also bends *oriT* DNA upon binding, these features together all form a highly organized conformation of *oriT* DNA-protein complexes. It has not been determined whether F TraM bends DNA upon binding. Electron microscopic analysis of pED208 TraM and *oriT* DNA failed to show bending of DNA (Di Laurenzio *et al.*, 1991). Recently, *in vivo* assays using KMnO4 to measure structural alterations at *oriT* showed that TraM failed to increase the sensitivity of *oriT* to this single stranded-specific agent (Deonier *et al.*, 1994). Preliminary evidence then suggests that TraM binding to *oriT* DNA does not severely alter the structure of the DNA.

IV. The role of TraY in *traM* expression

It is possible that nicking at *oriT* is an essential component of the signalling process, and that the nicking complex, or relaxosome, is required for TraM binding to *sbmC*. An interaction between the proteins involved in the nicking complex (TraI, TraY and IHF) and TraM could explain the plasmid specificities of TraM, TraI and TraY. If this were the case, this would change the previously predicted role of TraM in signalling (Willetts and Wilkins, 1984) from prior to the nicking reaction to subsequent to the the nicking reaction. This is supported by the results presented in Chapter 3, where constitutive *traY* expression was demonstrated to have a positive effect on *traM* expression. Not only is

very little TraM protein produced in the absence of TraY, but TraY appears to play a role in determining which TraM binding sites are occupied, demonstrated by the almost complete repression of Pm1 in pNY300 in the pOX*traY*244 mutant and increased expression from Pm2 (Figure 3.5). These results imply that TraM binds to *sbmA* preferentially under these conditions, a result confirmed by gel retardation assays (Chapter 5). A possible explanation for this effect could be that the binding of TraY to *sbyA* promotes binding of TraM to *sbmC*. Binding of TraM to *sbmC* would titrate out the limited amount of available TraM in the cell, releasing repression at Pm1, the major *traM* promoter. In the absence of TraY however, no binding of TraM to *sbmC* would occur, so all the available TraM would bind to *sbmA*, repressing expression from Pm1 completely. The increased expression from Pm2 seen under these conditions suggests that this promoter is not as important in TraM production as Pm1. In light of this model, one might expect that differential expression of *traM* from Pm1 and Pm2 would be observed in pOX*traY*244 in the absence of any high copy number plasmids expressing *traM*. However, mating assays suggest that a small amount of functional TraY may be expressed in this construct, as low level transfer is observed. This is particularly likely when one considers that the F *traY* gene appears to be a gene duplication, and the mutant was created by the insertion of a kanamycin resistance cassette into the latter half of the gene (K. Ippen-Ihler, personal communication). In the absence of multicopy plasmids bearing *oriT* then, enough residual TraY activity may remain to allow binding to the single *sbyA* site on pOX*traY*244. However, in the presence of mutiple *oriT*s bearing *sbyA*, this low level of truncated TraY may not be sufficient to bind all sites.

V. The role of TraM in transfer

Recently, an F plasmid relaxosome complex has been reconstituted *in vitro* (Matson *et al.*, 1994). Site- and strand-specific nicking of *oriT* DNA (linear or supercoiled) was detected in the presence of TraI, TraY and IHF, in a reaction that was absolutely dependent on the presence of all three proteins. Electron microscopic studies showed that a nucleoprotein complex containing all three proteins was formed, and footprinting studies demonstrated a region adjacent to *oriT* that was protected by TraI. It has previously been demonstrated that TraI complexes the 3' end of the nicked strand, such that the 3'-OH is only available for extension by DNA polymerase I by prior treatment with SDS and proteinase K (Matson *et al.*, 1993). In addition, the 5' end of the nicked strand is covalently attached to TraI. It has been demonstrated that TraI is capable of nicking and religating two single stranded oligonucleotides, implicating a role in the initiation and termination of strand transfer. Clearly, TraM has no role in the nicking of

oriT DNA, but perhaps it is required to alter the major activity of TraI from that of an endonuclease to that of a helicase.

In the F plasmid, the Py operon is constitutively expressed due to the absence of a functional FinOP fertility inhibition system. The presence of recipient cells would then presumably result in the formation of stable mating aggregates, and nicking would occur at *oriT* by the action of the relaxosome complex, consisting of TraI, TraY and IHF bound at their appropriate sites. In the absence of TraY (in a repressed cell), *sbmA* and *sbmB* would be occupied by TraM, repressing expression of *traM*. The binding of TraI-TraY-IHF complex at *oriT*, increases the efficiency of binding of TraM to *sbmC*, derepressing *traM* expression at Pm1. Interaction between TraM at *sbmC* and TraY at *sbyA* could be responsible for dissociating the interaction between TraI and TraY, invoking the helicase activity of TraI. The nicked strand could then be unwound and transferred to the recipient cell. This model also explains the requirement for *traM* in conjugal donor second strand synthesis (Kingsman and Willetts, 1978): unless TraI helicase activity generates a single strand for transfer, second strand synthesis will not occur. Increased *traM* expression as a result of TraY binding at *sbyA* will eventually lead to sufficient TraM in the cell to autoregulate Pm1 and Pm2, thus returning the system to a null state. A schematic diagram showing the proposed protein-DNA interactions in the absence and presence of TraY is shown in Figure 6.1A and B.

A role for TraM in transfer subsequent to *oriT* nicking leaves open the question of whether a specific signal is required to activate the nicking reaction in a repressed cell. Maximal expression of Py requires an upset in the fine balance between FinO, FinP antisense RNA, and *traJ* transcripts, such that *traJ* mRNA is translated and activates *tra* operon expression at Py. The regulation of Py is not understood, and the presence of a TraY binding site within the *traY* promoter region suggests that this gene may be autoregulated (Nelson *et al.*,1993). The *traY* promoter appears to be dependent on local superhelical density, which correlated well with the effect of TraJ on Py expression, suggesting that TraJ controls Py by affecting the superhelical density of the Py promoter (Gaudin and Silverman, 1993). However, a role for additional transcriptional activators in the expression of *traY* has not been ruled out.

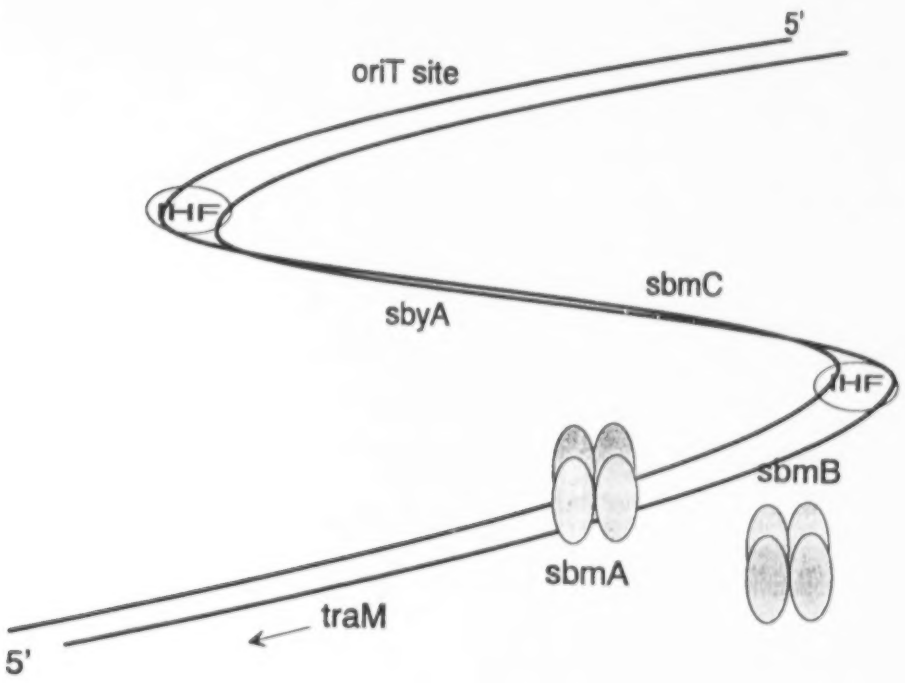
VI. Regulation of *traM* by FinO

The negative regulatory role of the *finO* gene product on *traM* expression was surprising, since this protein has never been implicated in transfer in a role other than that of

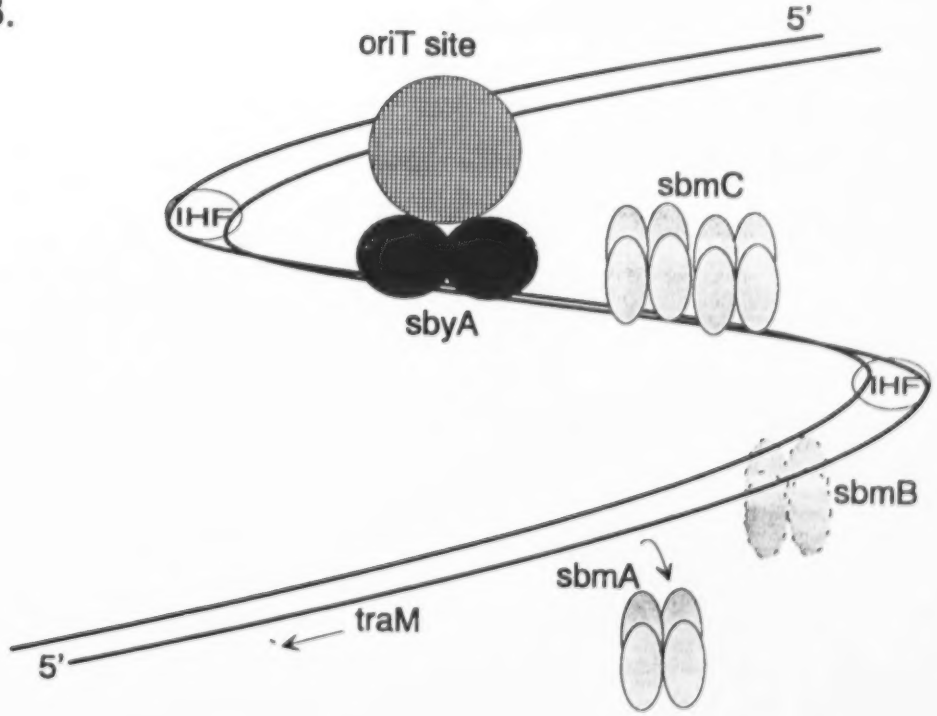
Figure 6.1

A model for the protein-DNA interactions at F *oriT* in the absence (A) or presence (B) of TraY protein. In the absence of TraY, TraM does not bind to *sbmC* and expression from Pm1 is repressed by the binding of TraM to *sbmA* (A). In the presence of TraY, *sbmC* is occupied by TraM and expression from Pm1 is derepressed (B). The conditions under which *sbmB* are occupied are presently unclear. It is proposed that in the presence of TraY, TraM interactions at *sbmC* promote the helicase activity of TraI, thus initiating transfer. For the sake of clarity, proteins are not shown to occupy their entire site on the DNA, but experimental evidence suggests that sites between *sbyA* and *sbmA* are immeadiately adjacent to each other. Also, there is no experimental evidence to suggest that two tetramers of TraM bind to *sbmC*, but is has been shown this way as the sequence at this site suggests that it binds TraM differently than *sbmA* and *sbmB*.

A.


5'
oriT site
IHF
sbmC
sbyA
IHF
sbmB
sbmA
traM
5'

B.


oriT site
5'
IHF
sbmC
sbyA
IHF
sbmB
sbmA
traM
5'

preventing *traJ* translation. A likely mechanism for this effect is that FinO, together with the FinP antisense RNA, prevents Py expression by inhibiting *traJ* translation. Results presented in Chapter 3 show that TraY is required for optimal TraM production, thus the presence of FinO in F-containing cells could inhibit the expression of *traM* through inhibition of *traY* expression. Confirmation of this requires the determination of whether the repression of *traM* expression by FinO can be detected in cells that do not express other transfer proteins, particularly TraY. An argument against this model is presented by the results of the analysis of TraM expression in the *traJ* mutant, F*lactraJ*90 (Figure 3.10). The model suggests that *traJ* would be expected to have an effect on *traM* expression, however, no evidence for this was found, either in work presented in this thesis (Chapter 3), or in the plasmids R1 (Schwab *et al.*, 1993), or R100 (Abo and Ohtsubo, 1993). It is possible that low level TraJ-independent expression of *traY* occurs in the *traJ* mutant, and that this provides sufficient TraY protein to derepress *traM* expression from Pm1. A suggestion that the *traJ*90 mutation is leaky (Achtman, 1973), and allows low level transfer, supports this possibility. Clearly, a better understanding of the regulation of *traY* expression is required before this question can be answered.

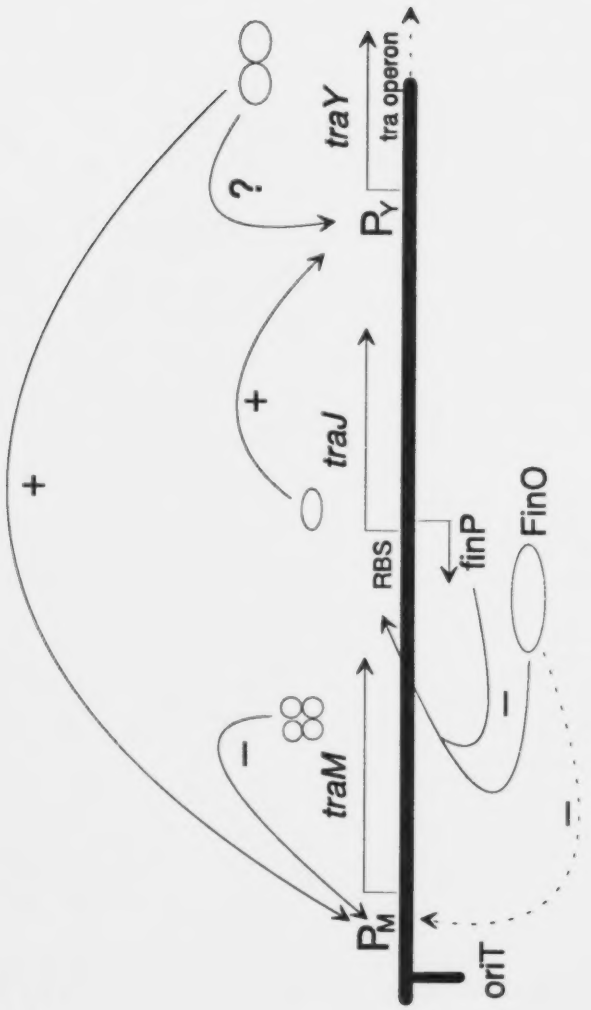
As described in Chapter 3, the regulation of conjugation is proving to be an extremely complex process, with host- and plasmid-encoded factors interacting at a number of levels to regulate transfer. The expression of *traM* has been studied in some detail in this thesis, and a summary of the control of *traM* expression is presented in Figure 6.2. Chapter 2 describes the identification of a rho-dependent terminator in the *traK* gene, which is recognised in the absence of translation and prevents unecessary transcription. Numerous features which fitted the criteria established for rho-dependent terminator consensus sequences were identified within the entire 33.3 kb transfer region, providing another example of regulation of gene expression from the F plasmid.

VII. Conjugation versus replication

The conservation of features within the *oriT* regions of F-like plasmids (Figure 5.5) and the general conservation of mechanisms of transfer in a wide variety of plasmids, is reminiscent of the conservation of structure of replication origins, described in Chapter 1. In particular, the major requirement for the initiation of replication is an unwinding of the duplex to allow the entry of the replication enzymes, specifically the DnaB helicase. In *oriC*, the chromosomal replication origin, this is accomplished by the binding of DnaA to specific recognition sequences within the origin. Theta-type plasmid replication requires plasmid specified proteins, the Rep proteins, in addition to DnaA to unwind the

Figure 6.2

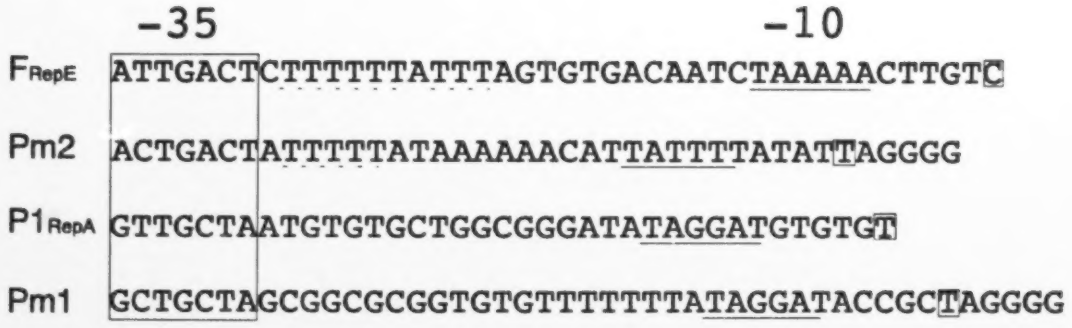
A summary of the regulation of *traM* expression. Pm represents the *traM* promoter region and Py represents the *traY* promoter. RBS is the ribosome binding site of the *traJ* gene, which is occluded by the action of the FinOP antisense negative regulatory system. '+' indicates positive regulation, while '-' indicates negative regulation of genes. The effect of the TraY protein on *traY* expression is not known and represented by a question mark.


+
?
traY
tra operon
P_Y
+
traJ
RBS
finP
FinO
–
traM
P_M
–
–
oriT

duplex. Interestingly, an analysis of the -10 and -35 regions of the *traM* promoter, Pm1, shows a remarkable similarity to that of the *repA* promoter from plasmid P1 (Figure 6.3). Pm2 bears some similarity to another *rep* promoter, the *repE* promoter of F, which has been shown to be controlled by σ^{32} -RNA polymerase. Another feature of σ^{32}-controlled promoters is the presence of a "T-tract" starting 4bp downstream of the -35 box, and this feature is also present in Pm2. Like the *traM* promoters, *rep* promoters are tightly regulated and the RepI protein of pColV-K30 has been shown to be expressed at very low levels (Perez-Casal *et al.*, 1989). In addition, like *traM*, many genes encoding these replication initiator proteins have been demonstrated to be autoregulated at the level of promoter repression (Gammie and Crosa, 1993). The model presented in this chapter suggests that TraM plays a role in inducing the helicase activity of TraI. It is tempting to speculate that TraM and TraY have similar functions in the regulation of transfer to those of DnaA and Rep proteins in replication. It remains to be determined whether the binding of TraY and TraM to the *oriT* region causes a structural change in the nucleoprotein complex which causes activation of the plasmid-encoded helicase, TraI, and thus initiates strand opening. Despite the apparent similarities of *traM* promoters to *rep* promoters, it should be recognised that mechanism of intiation of conjugation most strongly resembles the mechanism of initiation of rolling circle replication of Gram-positive plasmids such as pT181. In both systems, a site- and strand-specific nick is generated, followed by the unwinding of a single strand of DNA and subsequent second strand synthesis. One could suggest that the transfer mechanism of conjugative plasmids has evolved from a hybrid of theta- and rolling circle-replication mechanisms.

Figure 6.3

A comparison of the F *traM* promoter regions with those of the autoregulated *rep* promoters, RepE from the sex factor F and RepA from plasmid P1. The large box represents the -35 regions, while the -10 regions are underlined. The broken lines indicate T-tracts characteristic of promoters recognised by σ^{32}. Known +1 nucleotides are boxed.


-35
-10
FRepE ATTGACTCTTTTTTATTTAGTGTGACAATCTAAAAACTTGTC
Pm2 ACTGACTATTTTTATAAAAAACATTATTTTATATTAGGGG
P1RepA GTTGCTAATGTGTGCTGGCGGGATATAGGATGTGTGT
Pm1 GCTGCTAGCGGCGCGGTGTGTTTTTTTATAGGATACCGCTAGGGG

B. Future perspectives

With the recent reconstitution of a nicking complex of the F plasmid under conditions which mimic physiological conditions, we are in a position to clearly dissect the roles of the various proteins that interact at *oriT* to intiate transfer. Clearly, TraM is not involved in the nicking process, therefore another role for this protein must be found. Obvious questions that remain to be answered are whether *sbmC* is occupied *in vivo*, and whether additional *tra* proteins are required for binding of TraM to *sbmC*. *In vivo* footprinting experiments may be particularly useful in answering these questions, as *in vitro* results must be interpreted with some caution. Also, according to the model presented in this work, TraM and TraY interact to alter TraI activity. The addition of TraM to the nicking complex would be expected to alter the ATPase activity of the reaction, if this is true. In addition, the amount of TraM protein produced in F^+ cells in the presence of recipient cells has not been examined. According to the model presented here, a transient increase in TraM production would be expected under these conditions. In adddition, the role of FinO in *traM* expression needs to be examined more closely. The specificity of FinO binding to RNA has not been clearly elucidated, and would shed some light on whether the protein is capable of binding to *traM* mRNA.

The evidence presented in this thesis on the regulation of *traM* expression allows us to ask the pertinent questions on the cascade of events that lead to the conjugative transfer of DNA. By answering some of these questions, we will gain a better understanding of how this powerful mechanism of gene transfer is regulated and a better understanding of gene regulation in general.

C. References

Abo, T., and E. Ohtsubo. (1993) Repression of the *traM* gene of plasmid R100 by its own product and integration host factor at one of two promoters. *J. Bacteriol.* **175:** 4466-4474.

Achtman, M. (1973) Transfer-positive J-independent revertants of the F factor in *Escherichia coli* K12. *Genet. Res.* **21:** 67-77.

Achtman, M., and R. Skurray. (1977) A redefinition of the mating phenomenon in bacteria, p. 234-279. *In* Microbial Interactions, J. L. Reissig (ed.), vol. 3. Chapman and Hall, London.

Achtman, M., N. Willetts, and Clark, A.J. (1972) Conjugational complementational analysis of transfer-deficient mutants of F*lac* in *Escherichia coli*. *J Bacteriol* **110:** 831-842.

Dempsey, W.B., and Fee, B.E. (1990) Integration host factor affects the expression of two genes at the conjugal transfer origin of plasmid R100. *Mol Microbiol* **4:** 1019-1028.

Deonier, R. C., Y. Luo, Q. Gao, and A. Smorgorzewska. (1994) Genetic and structural analysis of the F plasmid *oriT* nicking region. (Abstract) Plasmid Biology. Banff, Canada.

Di Laurenzio, L. (1992) Protein-DNA interactions at *oriT* regions. Ph.D. Thesis. University of Alberta, Canada.

Di Laurenzio, L., Frost, L.S., Finlay, B.B., and Paranchych, W. (1991) Characterization of the *oriT* region of the IncFV plasmid pED208. *Mol Microbiol* **5:** 1779-1790.

Di Laurenzio, L., Frost, L.S., and Paranchych, W. (1992a) The TraM protein of the conjugative plasmid F binds to the origin of transfer of the F and ColE1 plasmids. *Mol Microbiol* **6:** 2951-2959.

Di Laurenzio, L., L. S. Frost, D. S. Scraba, and W. Paranchych. (1992b) Unpublished data.

Everett, R., and N. Willetts. (1980) Characterisation of an *in vivo* system for nicking at the origin of conjugal DNA transfer of the sex factor F. *J. Mol. Biol.* **136:** 129-150.

Gammie, A.E., and Crosa, J.H. (1991) Co-operative autoregulation of a replication protein gene. *Mol Microbiol* **5:** 3015-3023.

Gaudin, G. M., and P. M. Silverman. (1993) Contributions of promoter context and structure to regulated expression of the F plasmid *traY* promoter in *Escherichia coli* K-12. *Mol. Microbiol.* **8:** 335-342.

Kingsman, A., and N. Willetts. (1978) The requirements for conjugal DNA synthesis in the donor strain during F*lac* transfer. *J. Mol. Biol.* **122:** 287-300.

Luo, Y., Q. Gao, and R. Deonier. (1994) Mutational and physical analysis of F plasmid *traY* protein binding at *oriT*. *Mol. Microbiol.* **11:** 449-458.

Matson, S. W., M. T. Howard, W. C. Nelson, and J. A. Sherman. (1994) Characterization of the nicking and religation reactions catalyzed by DNA helicase I at the F plasmid origin of transfer. (Abstract). Plasmid Biology. Banff, Canada.

Matson, S. W., W. C. Nelson, and B. S. Morton. (1993) Characterization of the reaction product of the *oriT* nicking reaction catalyzed by *Escherichia coli* DNA helicase I. *J. Bacteriol.* **175:** 2599-2606.

Nelson, W. C., B. S. Morton, E. E. Lahue, and S. W. Matson. (1993) Characterization of the *Escherichia coli* F factor *traY* gene product and its binding sites. *J. Bacteriol.* **175:** 2221-2228.

Perez-Casal, J.F., Gammie, A.E., and Crosa, J.H. (1989) Nucleotide sequence analysis and expression of the minimum REPI replication origin and incompatibility deteminants of pColV-K30. *J Bacteriol* **171:** 2195-2201.

Schwab, M., H. Gruber, and G. Högenauer. (1991) The TraM protein of plasmid R1 is a DNA-binding protein. *Mol. Microbiol.* **5:** 493-446.

Schwab, M., H. Reisenzien, and G. Högenauer. (1993) TraM of plasmid R1 regulates its own expression. *Mol. Microbiol.* **7:** 795-803.

Tsai, M.-M., Fu, F.Y.-H., and Deonier, R.C. (1990) Intrinsic bends and integration host factor binding at F plasmid *oriT*. *J Bacteriol* **72:** 4603-4609.

Willetts, N. S., and B. Wilkins. (1984) Processing of plasmid DNA during bacterial conjugation. *Microbiol. Rev.* **48:** 24-41.

Yang, C.-C., and H. A. Nash. (1989) The interaction of *E. coli* IHF protein with its specific binding sites. *Cell* **57:** 869-880.